ARI Research Note 92-49

Handbook for Operating the OWLKNEST Technology

Regina M. Harris, Susan G. Hill, and Robert J. Lysaght
Analytics, Inc.

Richard E. Christ
U.S. Army Research Institute

Field Unit at Fort Bliss, Texas
Michael H. Strub, Chief

MANPRINT Division
Robin L. Keesee, Director

June 1992

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

MICHAEL D. SHALER
COL, AR
Commanding

Research accomplished under contract for the Department of the Army

Analytics, Inc.

Technical review by

Jon J. Fallesen
Rudy Laine
Linda G. Pierce
Bernard Schuster
Beverly Van Hoff

Accession For	
NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1992, June	Interim Sep 86 - Jul 90

4. TITLE AND SUBTITLE

Handbook for Operating the OWLKNEST Technology

5. FUNDING NUMBERS

MDA903-86-C-0384
62785A
790
1102 (113)
C01

6. AUTHOR(S)

Harris, Regina M.; Hill, Susan G.; and Lysaght, Robert J. (Analytics, Inc.); Christ, Richard E.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Analytics, Inc.
2500 Maryland Road
Willow Grove, PA 19090

8. PERFORMING ORGANIZATION REPORT NUMBER

--

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Institute for the Behavioral and Social Sciences
ATTN: PERI-S
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

ARI Research Note 92-49

11. SUPPLEMENTARY NOTES

Contracting Officer's Representative, Richard E. Christ. Other authors are currently with the following organizations: R. M. Harris--Hilton Systems, Inc.; S. G. Hill--EG&G Idaho, Inc.; and R. J. Lysaght--Bellcore.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release;
distribution is unlimited.

12b. DISTRIBUTION CODE

--

13. ABSTRACT *(Maximum 200 words)*

This handbook offers guidance to users of the operator workload (OWL) knowledge-based expert system tool (OWLKNEST). The handbook is an appropriate foundation for understanding the context within which OWLKNEST can be useful. The handbook describes (a) the concept of workload; (b) problems associated with determining which OWL assessment techniques are most appropriate for any given set of study objectives, system characteristics, and user resource constraints; and (c) the expert system approach to problem solving. The handbook also provides an overview of OWLKNEST and gives detailed instructions and specific examples for operating the tool and interpreting its output. Brief descriptions of each of the 38 OWL assessment techniques included in OWLKNEST and a survey form designed to solicit feedback from OWLKNEST users are in appendixes.

14. SUBJECT TERMS

Operator workload	Expert system	Workload assessment
Measurement	Human performance	Workload prediction
MANPRINT	System development	

15. NUMBER OF PAGES

158

16. PRICE CODE

--

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

FOREWORD

The MANPRINT Division of the U.S. Army Research Institute for the Behavioral and Social Sciences supports the Army with research and development on manpower, personnel, training, and human performance issues as they affect the development, acquisition, and operational performance of Army systems and the combat readiness and effectiveness of Army units. One concern that underlies all these issues is the mental workload imposed on the operators of newly emerging, high-technology systems and the impact of that workload on operator and system performance. The Fort Bliss Field Unit is conducting exploratory development research to establish the foundation for an operator workload (OWL) assessment program for the Army.

This handbook and the accompanying computer software, the OWL knowledge-based expert system tool (OWLKNEST), are major outputs of the OWL assessment program. Together they aid military analysts by providing recommendations for workload measurement techniques that are most appropriate for any given set of workload study objectives, system characteristics, and available user resources. The handbook also illustrates how OWLKNEST can be used to gain insights on the appropriateness of various assessment techniques at different stages in the system development cycle and to perform sensitivity or what-if analyses in which the user changes one or more responses to the questions posed by the expert system.

ACKNOWLEDGMENTS

This report benefited from the significant contributions of many individuals. First and foremost, special thanks is due to those listed as technical reviewers. Each reviewed the handbook and offered suggestions for improving not only the handbook but also the OWLKNEST software itself.

Appreciation is also extended to Harold Booher of the Directorate for MANPRINT in the Office of the Deputy Chief of Staff for Personnel at Headquarters, Department of the Army, and to Ray Brandenburg of the MANPRINT Programs Division of the U.S. Army Personnel Integration Command. These individuals and their staffs have contributed general and continuous guidance, inspiration, and support to this specific product as well as to the overall operator workload research program at the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI).

The authors acknowledge the contribution of other members of the original OWL team, including Helene P. Iavecchia, Alvah C. Bittner, Jr., A. O. Dick, Paul M. Linton, James C. Byers, and Brian D. Plamondon, and also Michelle R. Sams of the ARI Fort Bliss Field Unit. Each of them participated in activities essential to the overall research task, which provided the knowledge and experiences necessary for preparing this report. Each also contributed thoughts on its form and substance. The production of final drafts of this handbook and its accompanying software was performed using facilities made available by Hilton Systems, Inc., and Chi Systems, Inc.

HANDBOOK FOR OPERATING THE OWLKNEST TECHNOLOGY

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

HANDBOOK FOR OPERATING THE OWLKNEST TECHNOLOGY

PART ONE: Introduction

Purpose

The Operator Workload Knowledge-based Expert System Tool (OWLKNEST) is a microcomputer-based tool that provides guidance in selecting the most appropriate technique to use for assessing Operator Workload (OWL) during the system acquisition process. This Handbook for Operating the OWLKNEST Technology (HOOT) is a user's guide for Version 1 of OWLKNEST (released in February, 1991). It describes the underlying rationale, capabilities, and features of OWLKNEST in addition to giving instructions for using the tool.

Overview of HOOT

The remainder of the handbook is organized into the following parts:

PART TWO: Operator Workload

Provides an overview of operator workload (OWL) and of techniques used to assess OWL. Also describes the research program which led to the development of OWLKNEST.

PART THREE: EXPERT SYSTEMS

Provides a brief description and discussion of expert systems along with an overview of the major components of an expert system.

PART FOUR: OWLKNEST OVERVIEW

Contains an overview of OWLKNEST. Presents information about the knowledge included in OWLKNEST and how the knowledge is organized. Describes how OWLKNEST is used to define the parameters for a particular workload assessment and to obtain recommendations.

PART FIVE: Installing OWLKNEST

Describes the hardware and software environment required for OWLKNEST and the procedure for installing OWLKNEST on your microcomputer.

PART SIX: OPERATING OWLKNEST

Describes how to start the OWLKNEST software and explains the components of the OWLKNEST user-computer interface and how to use them.

PART SEVEN: SAMPLE PROBLEMS

Presents examples of using OWLKNEST for specific problems. Describes how to use the OWLKNEST recommendations and strategies for handling too few or too many recommendations.

PART EIGHT: CONCLUSIONS

Summarizes the benefits of utilizing OWLKNEST and describes the survey form.

Parts Two and Three may be treated as independent sections of this manual. Hence, if the reader has no need to review issues related to workload or to review the major components of an expert system, Parts Two or Three, respectively, may be skipped. However, the OWLKNEST overview (in Part Four), installation and operation procedures (in Parts Five and Six), and example applications (in Part Seven) should be read carefully, the latter three while simultaneously interacting with OWLKNEST as the software program is loaded and operated on a personal computer.

User Characteristics

The target user population for OWLKNEST and HOOT are the analysts involved in assessing operator workload for a military (or commercial) system. Potential users are Army MANPRINT analysts or human factors specialists. This assumes that users have at least a fundamental knowledge of workload and human performance concepts and that the users possess basic knowledge about how to operate a computer, such as how to turn on the computer and insert a floppy diskette, but does not assume any knowledge of expert systems.

PART TWO: Operator Workload

Problem

Projected manpower declines coupled with increases in personnel costs and battlefield sophistication has prompted an increased reliance on high technology equipment in new Army systems. As technology has changed, the role of the operator has also changed. Task requirements for the operator have shifted from those that primarily require physical exertion to those which have increasingly larger amounts of perceptual and cognitive demand.

While technological advancements may increase system capability, it is critical to ensure that the resulting systems do not concurrently cause the demand for mental skills to exceed the operator's capabilities. Task demands greater than an operator's capacity to respond may result in undesirable consequences, such as mission degradation or failure, or compromised system safety.

Operator Workload Definition

The concept of work in the physical sciences is readily understood; work is not performed without some expenditure of energy or other resources, and work rate or work efficiency may change depending on the demands of the situation. Likewise for the human, both physical and mental work depend not only on the particular task to be accomplished, but also upon the availability of the internal resources required of the operator to perform the task. Thus, operator workload (OWL) is generally defined in terms of the interaction between the work imposed on an operator by a task and the operator's capacity to perform that work. The conceptual foundations of workload have been adequately discussed in numerous recent publications (e.g., Gopher & Donchin, 1986; Lysaght et al., 1989).

Operator Workload Assessment Techniques

A variety of OWL assessment techniques are available and many have been documented in published papers (e.g., Lysaght et al., 1989; O'Donnell & Eggemeier, 1986; Wierwille & Williges, 1980). Workload assessment methods include *analytical* or predictive techniques which may be applied early in system design without an operator "in-the loop" and *empirical* techniques which require an operator using a simulator, prototype, or representative system. Analytical techniques are used to predict performance and estimate workload through the methods of expert opinion, comparability analysis, task analysis, and simulation models. Empirical techniques include methods which measure the operator's performance, subjective experience, and physiological responses.

Operator workload analysts have found it difficult to readily determine which technique is most appropriate for their particular workload study (Hill & Harris, 1989). Aside from a large number of assessment methods from which to choose, the analyst must also consider the objectives of the workload study in relation to the characteristics of candidate techniques. For example, workload assessment techniques differ in their sensitivity and diagnosticity. Sensitivity refers to the degree to which the technique can differentiate between different levels of workload experienced by the operator as affected by distinct levels of task demands. Diagnosticity refers to the extent to which a technique reveals not only the overall level of OWL but also information about the component factors that contribute to overall OWL (e.g., perceptual, cognitive, and psychomotor factors). The selection of the optimum technique is further complicated by real world constraints (e.g., time, cost, personnel requirements, and facilities).

Based on information gathered from Army personnel and documents, it is evident there is a void in specific guidance concerning the implementation of operator workload assessment during the Army materiel acquisition process. Developers of Army systems are required to conduct

workload analyses during the system development process under the general requirement of MIL-H-46855B and the specific mandate of AR 602-2 (MANPRINT) (see Christ, Bulger, Hill, & Zaklad, 1990, and Hill et al., 1987, for a discussion of U.S. Army operator workload assessment requirements). However, these requirements do not provide any guidance about how to perform the workload assessment or which techniques are to be used.

Operator Workload Program

In response to the need for useful guidance in the assessment of operator workload, the U.S. Army Research Institute (ARI) sponsored a three-year exploratory development research effort called the Operator Workload or OWL Program. The OWL program was to establish guidance for the assessment of operator workload associated with the operation of Army systems. In order to accomplish this objective, Army needs were first identified (Hill et al., 1987) and OWL assessment techniques were critically and comprehensively reviewed (Lysaght et al., 1989). Following this fundamental research, OWL field assessment and validation efforts were planned and conducted on three diverse Army systems at various stages of development and fielding:

- Aquila remotely-piloted vehicle (see Byers, Hill, Zaklad, & Christ, 1990),
- Line of Sight-Forward (Heavy) air defense system (see Hill, Byers, Christ, & Zaklad, 1989), and
- UH-60A BLACK HAWK helicopter (see Iavecchia, Linton, Byers, & Harris, 1989).

The results of these efforts under the OWL Program were the foundation for the development of three research products:

1. A pamphlet for Army managers that describes the need and some procedures for incorporating OWL issues and concerns into the Army material acquisition process (Christ, Bulger, Hill, & Zaklad, 1990),

2. A technical report that describes the empirical validation and application of some OWL assessment techniques having high potential utility in operational environments (Hill, Iavecchia, Zaklad, Christ, & Sams (1991), and
3. A computer-based tool, described in this handbook, that provides guidance in selecting OWL assessment techniques.

Approach to Providing Guidance for OWL Assessment

Rather than having yet another written manual (with its inherent difficulties associated with revisions and usability), the Army community expressed a desire for a computer-based guidance tool (Hill et al., 1987). As such, one of the products of the OWL Program is the Operator Workload Knowledge-based Expert System Tool (OWLKNEST), an interactive, computerized decision aid. As well as providing recommendations for workload assessment techniques, OWLKNEST is envisioned to serve as a clearinghouse of knowledge for workload assessment methodologies.

The next two parts of this handbook provide an introduction to expert systems and a description of the expert system approach used by OWLKNEST to providing guidance for selecting OWL assessment techniques, respectively.

PART THREE: Expert Systems

What is an Expert System?

According to Fotta and Davis (1988) and Miller (1988), an expert system is a computer program used to solve problems that are difficult enough to require significant human expertise for their solution. The program encodes human expert knowledge and reasoning processes from a specific, limited field (called the domain of the system). This expertise is stored in a highly usable, easy-to-access manner. Specifically, an expert system applies rules to determine the selection and presentation of questions posed to a user and, depending upon the answers provided by the user, also to determine the selection and ratings of the recommended solutions to the problems. Consequently, someone who is not an expert should be able to access and apply the encoded expertise to solve a problem from the domain of the system.

How Does an Expert System Work?

In general, two kinds of information are needed to solve any particular problem. The problem solver must have

1. A background of in-depth information or *expertise* in the domain of the problem, and
2. Specific information or *data* pertaining to a particular instance of the problem.

An example taken from Miller (1988) illustrates both the problem solving process and the applicability of expert systems to that process. Assume you have a problem with your automobile. Not being an expert in this problem domain, you take it to an automobile mechanic. To fix your car the mechanic needs to know a lot about cars (i.e., have expertise) and needs to know specifically what your car is doing or not doing (i.e., have

data). If you tell the mechanic about your car's specific problem, a good mechanic can figure out what needs fixing and how to implement the fix.

Now assume you do not have access to a mechanic but you have an expert system that contains the mechanic's expertise. If you feed into the expert system the data about your car's particular problem, the expert system can determine what it will take to solve that problem. (While a good mechanic can actually fix the car, the expert system cannot. However, a good expert system can tell you how to fix it).

An expert system, then, works by applying expertise to data. The data are provided by the user of the expert system and the expertise, which is encoded in the expert system, is supplied by a human expert. Also programmed into the expert system is a mechanism (called an inference engine) for applying the expertise to the data. The inference engine gives the expert system the power to reason and make decisions. The various components of an expert system are illustrated in Figure 3-1 and described in the next section.

Basic Components of an Expert System

Domain expert(s)

An individual or group of individuals who are widely recognized as being able to solve a particular type of problem more efficiently and effectively than most other people.

Knowledge engineer

A specially trained individual who interacts in numerous ways with the expert(s) to understand the essence of the expertise, describe the essential facts and the rules that operate on those facts, and encode the expertise into the knowledge base of an expert system.

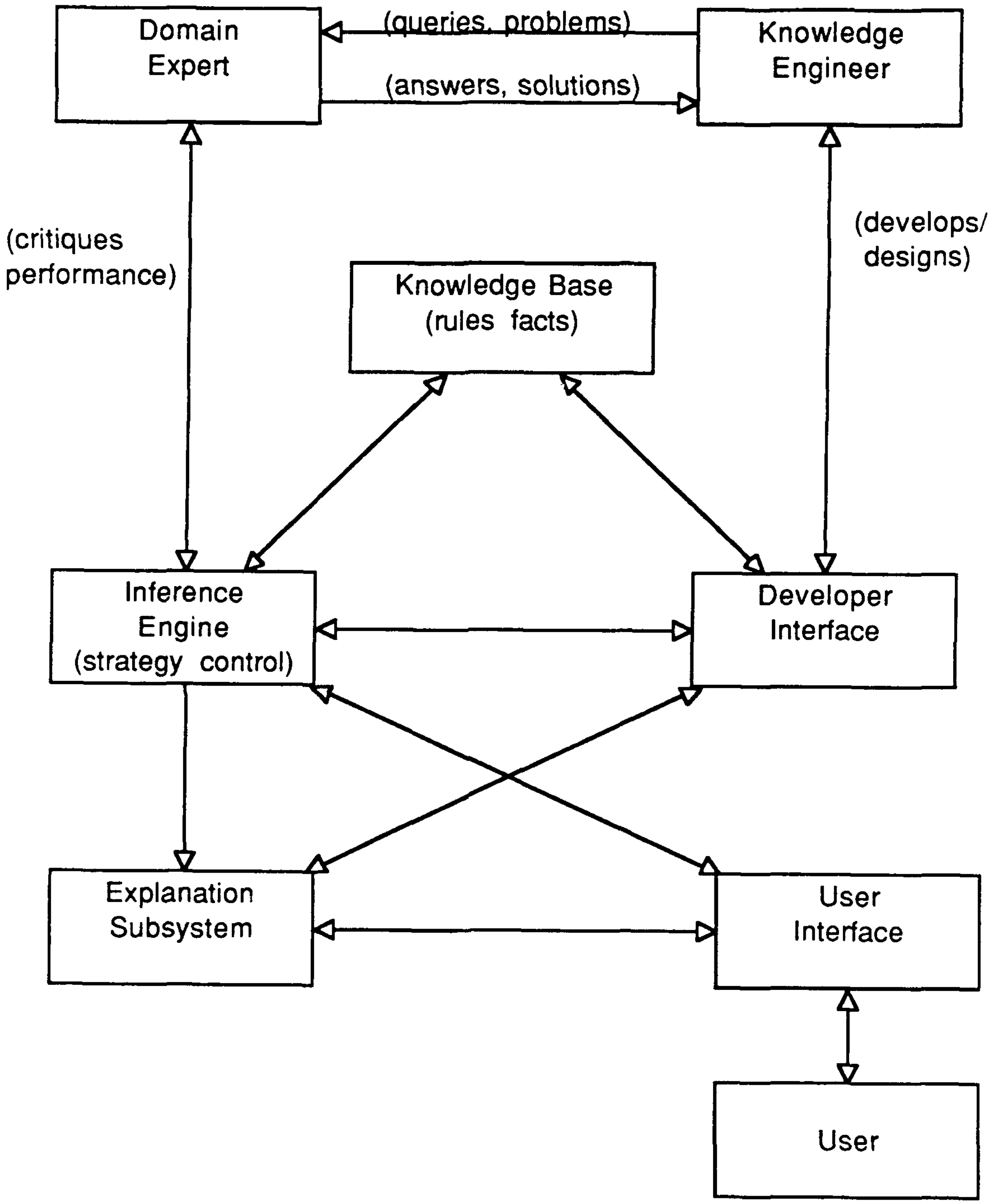


Figure 3-1. Basic Components of an Expert System

Knowledge base

The knowledge base of an expert system is comprised of both descriptive and procedural knowledge. Descriptive knowledge, also called assertions or facts, consists of values assigned to the attributes or features of objects, events, or ideas. These values are required to discriminate among the elements in the knowledge domain. For example, in the large domain of biology, facts consist of the following:

- A bird is an animal.
- A male cardinal is red.

In the domain of workload assessment techniques, facts may be expressed as relative rather than absolute values. Hence, two or more elements belonging to the domain are compared using terms such as less than, equal, or greater than. The following is a fact in the domain of workload assessment techniques.

- It requires more time for an OWL analyst to prepare the necessary detailed methods for a structured interview than for an unstructured interview.

Procedural knowledge consists of rules that address the relationships among facts or elements of descriptive knowledge in the knowledge domain. These rules normally are expressed as follows: "If ... (a set of premises are true), then ... (perform a specified action or make a specific conclusion)." Generally, both the premise and the conclusion are statements of descriptive knowledge, consisting of values assigned to attributes of elements in the knowledge domain. In the domain of biology, the following rules exist:

- If the animal is a bird, then the animal has feathers.
- If the animal is a bird, then the animal has wings.

In the domain of workload assessment techniques, the following rules are generally accepted:

- If the time to prepare detailed methods for an OWL assessment is short, then give a lower recommendation to the technique of structured interview than unstructured interview.
- If the time to prepare detailed methods for an OWL assessment is sufficiently short, then eliminate the technique of structured interview from further consideration.

Inference engine

An inference engine is a computer program that gives the expert system the capability to select and apply a sequence of rule-encoded expertise to the data of a specific problem. This capability essentially gives the expert system the power to reason and draw conclusions. The inference engine can accomplish this feat because it encodes strategies for problem solving that are borrowed from formal logic and search algorithms.

An inference engine enables the expert system to evaluate existing facts and rules. The inference engine also determines where to start a sequence of inferences, the order in which facts and rules are examined, and which type or line of reasoning is to be followed.

In general, the reasoning process in the inference engine is the path the computer follows as it traces rules through a knowledge base. The inference engine can employ both forward and backward reasoning or chaining procedures as appropriate. Forward chaining is a data-driven search procedure which matches the input data against the IF-part of a rule to formulate new facts from the THEN-part of the rule. In backward chaining the logic is traced from the conclusion back to the facts upon which it depends. Backward chaining procedures attempt to prove a new rule by determining what data are required to make the premise of a rule true.

Explanation subsystem

In addition to possessing practical knowledge that can be used to solve a particular type of problem, a human expert should also be able to explain the reasoning process that leads to a specific conclusion. The explanation subsystem of an expert system is designed to provide a similar

service to the user — if the user wants this information. Indeed, a non-expert may call upon a human expert or use an expert system to solve problems without caring about or wanting explanations. The reasoning process is normally invisible to the non-expert.

However, for a number of reasons, the user may wish to understand the problem-solving process. While an expert system cannot explain what it does in the same way a human expert can, it can answer the two questions of why and how. The user may ask *why* questions when the expert system requests data to match against the premise of rules. Essentially, the user asks the expert system why it wants a given type of data. The explanation subsystem can display the rule which uses the requested information to reach a given conclusion.

The user may ask *how* questions when the expert system reports the results of its reasoning about goals and subgoals, i.e., when it reports the rule conclusions. Essentially, the user asks the expert system how it arrived at a particular conclusion. The explanation subsystem will display the rules used.

User

The user of an expert system is an individual with a specific problem who also processes a working memory from which values or data pertaining to that problem may be elicited for input to the expert system. Except for only very basic knowledge about how to operate a computer, the user need not have any knowledge about how computers or expert systems work.

User interface with the expert system

One very important component of an expert system is the part of the program that allows the user to input the specifics of the current problem. This component is called the user interface. A good or "friendly" interface asks questions that the user can easily answer. A friendly interface presents a menu of alternative answers for each question so the user knows what

data are acceptable. Other favorable features of a good interface include a plausible sequence of questions, consistent prompts, a capability to review previous inputs, and an explanation subsystem.

Developer interface

Another important component of an expert system is the part of the program that enables the domain expert or knowledge engineer to encode the domain-specific knowledge base and to refine certain aspects of the inference engine. The criteria for evaluating the friendliness of the expert system's interface with its developer include the requirements and provisions for entering the expertise, text editing features, debugging aids, and external file support.

Expert system shell

Expert system developers generally separate the rules of a particular knowledge base from the inference engine used to process those rules. It can be shown that the inference engine or reasoning capability of different expert systems are essentially the same. The difference among expert systems is due to the different specific rules generated by the domain of each system.

The inference engine without any domain-specific knowledge base is called an expert system shell. A good expert system shell will accommodate different sets of domain-specific rules to create different expert systems. A useful shell also must have friendly interfaces with the developer and the user of the expert system.

Summary

While this handbook and the OWLKNEST technology do not require that the user have any knowledge about expert systems, an overview of the purpose and components of expert systems is provided for interested users. If this brief overview has whetted the readers' appetite for more information about expert systems, they are referred to Harmon and King

(1985), Waterman (1986), Boehm-Davis (1988), and Ignizio (1990) for a more thorough introduction. Boehm-Davis edited a special issue on Expert Systems for the Journal of the Human Factors Society. Ignizio is the author of a book that explains the fundamentals of rule-based expert systems which use the Exsys expert system shell employed for OWLKNEST.

PART FOUR: OWLKNEST Overview

This part of HOOT presents an overview of OWLKNEST, its knowledge base, expert system shell, and output. It also contains a brief description of the expert system predecessor to OWLKNEST.

OWLKNEST Description

OWLKNEST is based upon an expert system approach to problem solving. The expert system approach has been found to be particularly successful for classification-type applications which recommend an answer or solution to a problem from a set of alternatives based upon user inputs.

The goal of OWLKNEST is to provide the user with a prioritized list of operator workload techniques that best meet the needs of a particular workload study. In order to characterize the key features of the workload study, OWLKNEST poses a series of questions to the user regarding the objectives of the workload study and the resources available to conduct the study. The possible answers to each question are listed as numbered options. The user selects responses by entering the option number.

Figure 4-1 illustrates the OWLKNEST flow of information. The user inputs are fed into the expert system, which applies rules and knowledge dependent on the responses supplied by the user. After responding to all the questions, a list of the recommended workload techniques is displayed along with their associated numeric rating values. These values indicate the relative appropriateness of each technique for the particular situation described by the input of the user. At any point during the session, the user may ask for help from OWLKNEST to clarify a question, explain the current options, or show which rules are being applied.

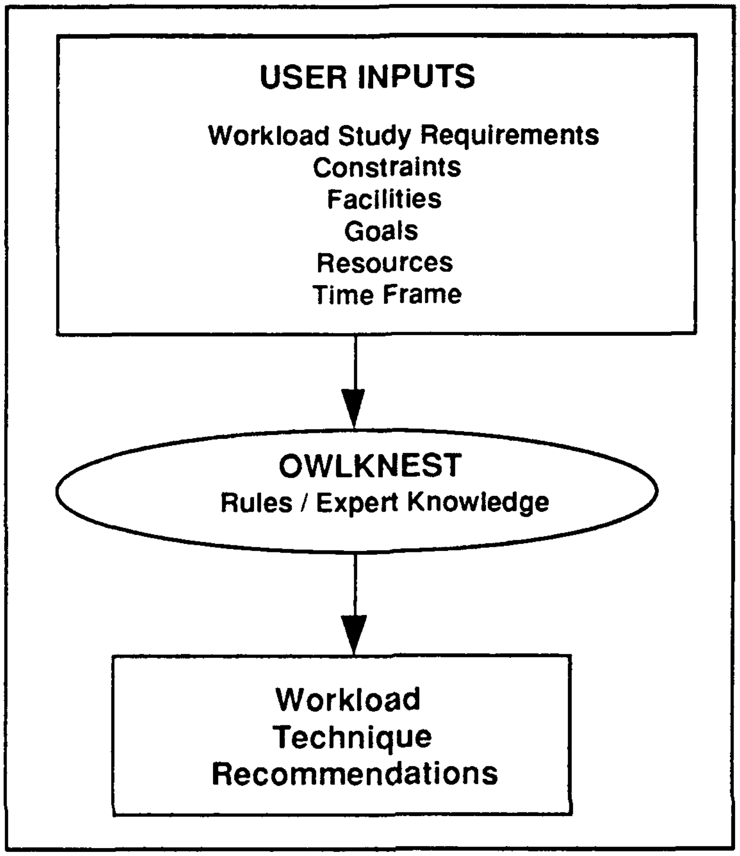


Figure 4-1. OWLKNEST Information Flow

OWLKNEST Knowledge Base

The OWLKNEST knowledge base contains information needed to differentiate among workload assessment techniques. Succeeding paragraphs of this section describe the source of the domain expertise, the OWL assessment techniques that are the foundation of the knowledge base, the structure of the knowledge, the salient features of the techniques, and the rules that are used to select questions for the user of OWLKNEST or to assign values to the assessment techniques.

Source of domain expertise

OWLKNEST incorporates a rule-based knowledge representation system. The domain experts responsible for developing the rules used to form this knowledge base were members of a team of human factors specialists assigned to the ARI OWL research program. Collectively, these individuals had over 70 years of experience in assessing workload and in utilizing knowledge gained from human factors studies conducted in military settings.

The primary knowledge source for OWLKNEST was derived from the comprehensive review and evaluation of current OWL assessment techniques (Lysaght et al., 1989), discussions with workload experts, and personal field experiences in utilizing various workload assessment techniques. What specific facts and rules were to be included in the knowledge base and the exact form of the reasoning process that was to be used to process that knowledge were determined by a consensus of the team's judgments.

Workload assessment techniques

The OWLKNEST knowledge base is built upon information about workload assessment techniques. Although numerous workload assessment techniques have been developed, their applicability to Army systems, usability, and capabilities vary greatly. Therefore, a set of evaluation criteria was developed to quantify the various techniques and to determine those suitable for inclusion in OWLKNEST.

Based upon the OWL team's judgement on the robustness of these techniques as applied to Army systems, a core set of techniques was identified which met the following evaluation criteria:

1. Demonstrated efficiency in application to real-world problems of Army systems,
2. Sufficient body of documentation not only on previous application of the technique but also on how to use the technique,

3. Availability in the Army and public domain, and
4. Sufficient evidence of validity.

Sixteen analytical or predictive OWL assessment techniques and twenty-two empirical or evaluative OWL assessment techniques met these criteria and are included in the OWLKNEST knowledge base. These two sets of techniques are listed in alphabetical order in Table 4-1.

It must be noted that some techniques that might seem on the surface to have applicability to this knowledge base are not included since there is no documented evidence that they have ever been successfully used to assess OWL. For example, there is one documented attempt to use comparability analysis techniques for predicting OWL (Shaffer, Shafer, & Kutch, 1986). However, neither the Comparison-Based Prediction (CBP) technique (John, Klein, & Taylor, 1986) nor the Early Comparability Analysis (ECA) technique (U.S. Army, 1987) have been used specifically to assess OWL. Hence, while the general case of comparison techniques is included in the OWLKNEST knowledge base, more specific examples of that set of techniques are not included.

Structural hierarchy of assessment techniques

The OWLKNEST knowledge base is organized according to the taxonomy suggested by Lysaght et al. (1989) which divides OWL techniques into analytical (predictive) and empirical (evaluative) techniques. Within each of these two divisions, other categories of workload techniques were identified. These workload techniques were organized into a classification (or decision) tree that resulted in an hierarchical knowledge structure. The upper part of the classification tree is illustrated in Figure 4-2. The major nodes shown are expanded as described below:

- Figure 4-3 shows the full tree for the node that represents the analytical expert opinion techniques.
- Figure 4-4 shows the expanded decision tree for the analytical simulation and task analysis techniques.

Table 4-1. OWLKNEST Operator Workload Techniques

Analytical
Closed Questionnaires
Comparability Analysis
Delphi Interviews
Human Operator Simulator
McCracken-Aldrich Task Analysis
MicroSaint
Open Ended Questionnaires
Prospective OW
Prospective SWAT
Prospective TLX
SIMWAM
Structured Interviews
TAWL
Tr/Ta Task Analysis
Unstructured Interviews
Zaklad/Zachary Task Analysis

Empirical
AHP
Bedford
Blink Rate
Choice RT Secondary Tasks
Closed Questionnaires
Embedded Secondary Tasks
Evoked Potentials
Eye Movement
Heart Rate
Heart Rate Variability
Modified Cooper-Harper
NASA-TLX
Open Ended Questionnaires
OW
Pupil Measures
Steinberg Memory Secondary Tasks
Structured Interviews
SWAT
Time Estimation Secondary Tasks
Type 1 Primary Measures
Type 2 Primary Measures
Unstructured Interviews

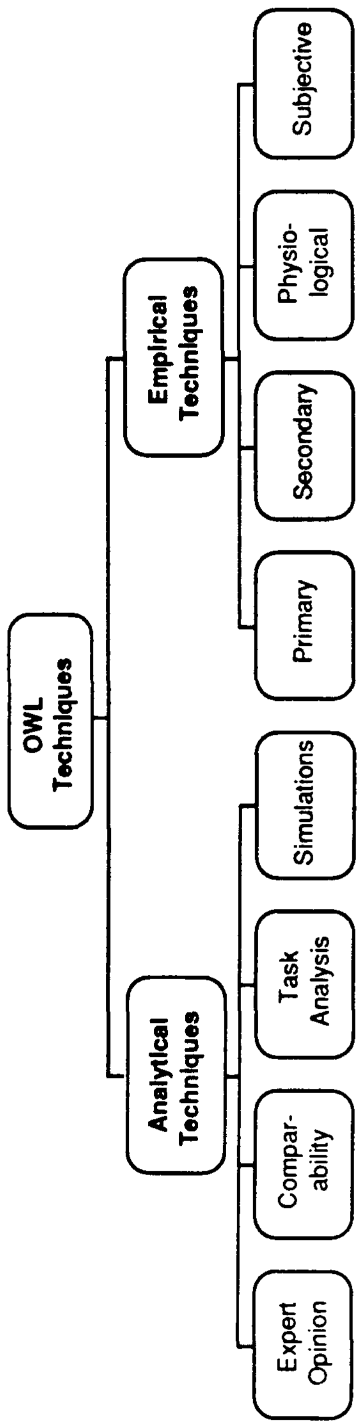


Figure 4-2. OWL Measurement Techniques Classification Tree

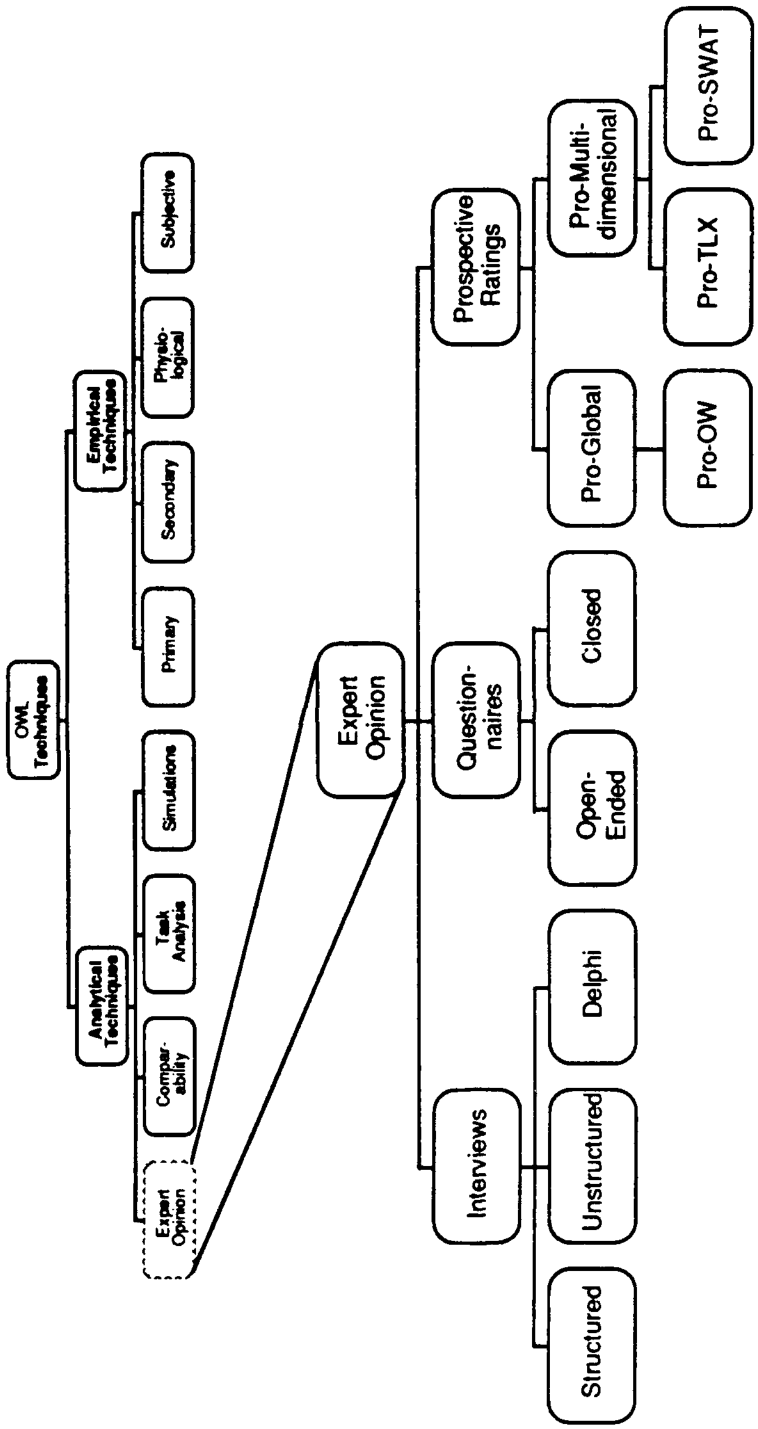


Figure 4-3. OWL Expert Opinion Techniques Classification Tree

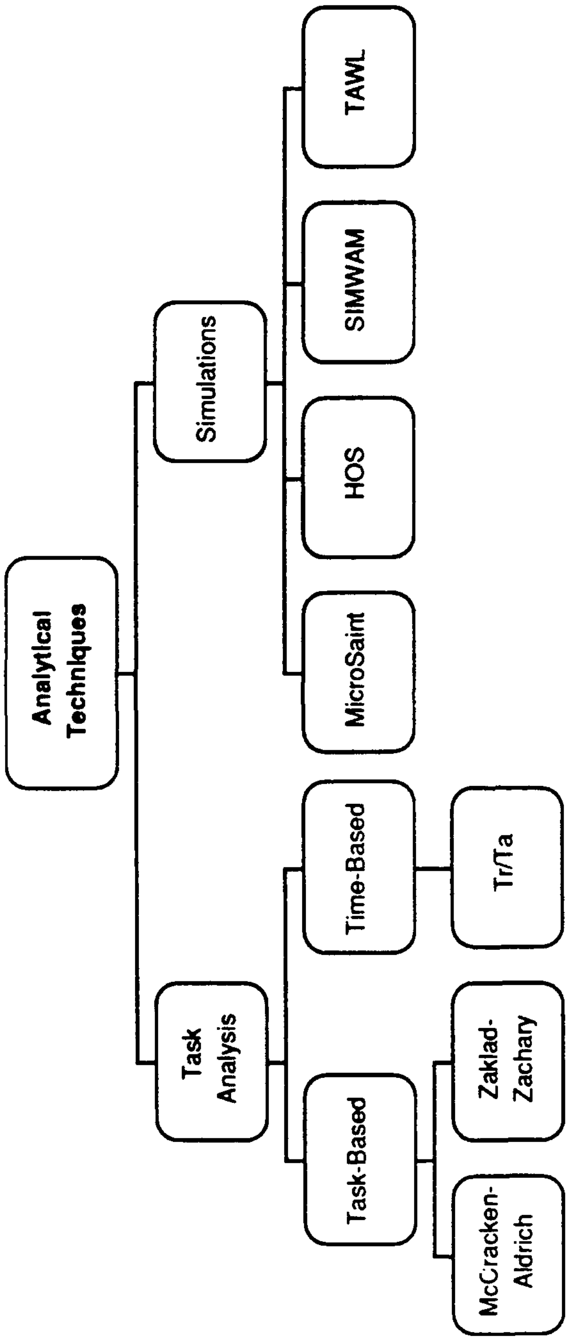


Figure 4-4. OWL Task Analysis and Simulation Techniques Classification Tree

- Figure 4-5 shows the expanded decision tree for the empirical primary and secondary techniques.
- Figure 4-6 shows the expanded decision tree for the empirical physiological techniques.
- Figure 4-7 shows the expanded decision tree for the empirical subjective techniques.

Some higher level nodes, such as that representing Comparability Analysis, represent generic classification of techniques that are not further subdivided. The terminal nodes of the decision tree structure represent the actual workload techniques as listed in Table 4-1. This structure provides a framework in which additional knowledge can be readily incorporated — new nodes can be easily incorporated into the tree or entire branches modified or deleted.

Salient features of the techniques

The key step in the development of the knowledge base was determining the key features or criteria that distinguish a given technique from others and determine the most suitable type of application. Some of the criteria are based on specific requirements (e.g., requires an IBM PC microcomputer) while others are less tangible (e.g., "easy-to-use"). Each of the key features results in a set of questions that will be posed to the user by the expert system. The answers to these questions will be used by the expert system to determine what techniques fit the user's workload study requirements as well as the user's resources. An important goal in selecting these features was to minimize the number of questions by selecting only those that were necessary and resulted in clear distinctions among the techniques.

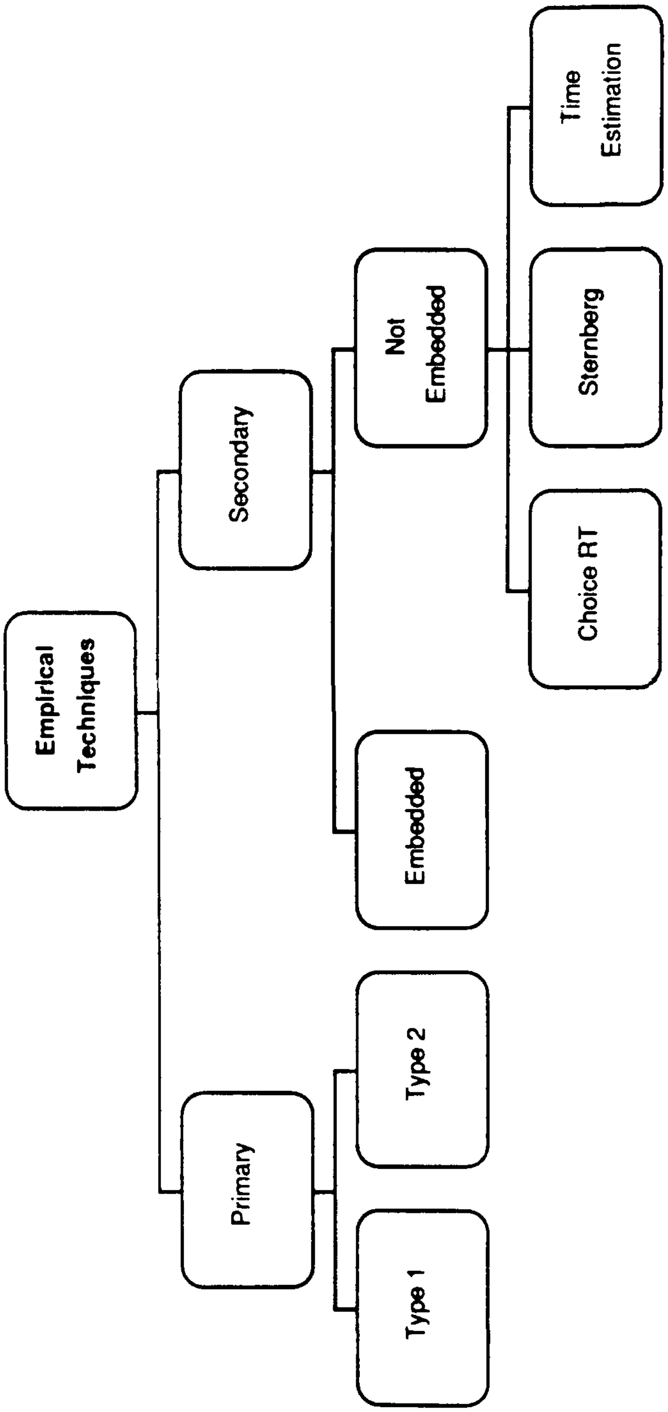


Figure 4-5. OWL Primary and Secondary Techniques Classification Tree

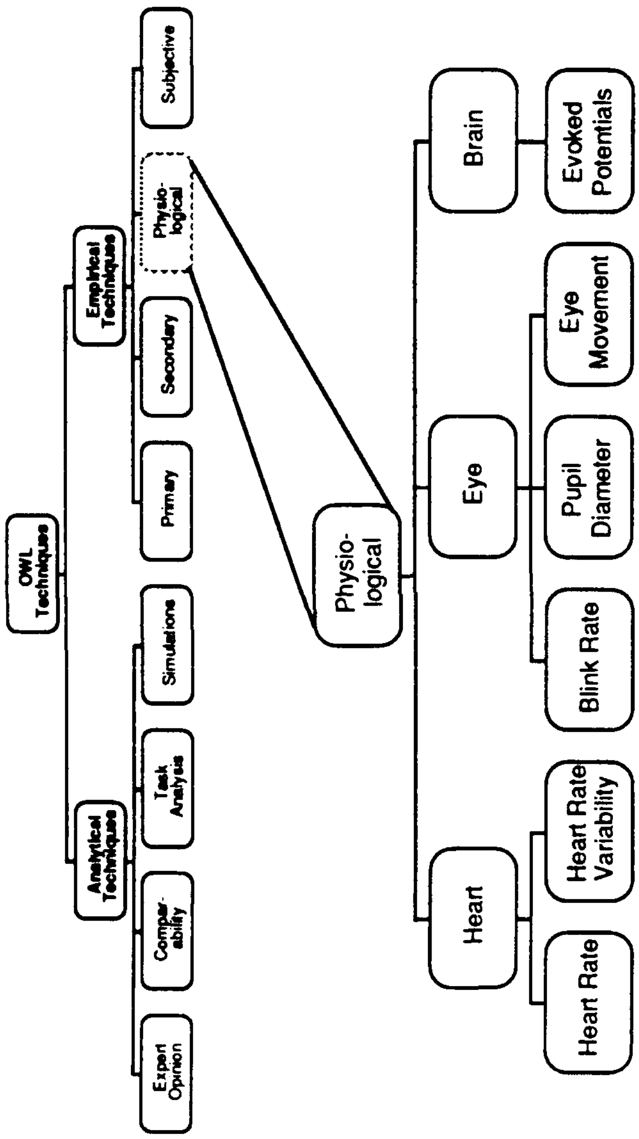


Figure 4-6. OWL Physiological Techniques Classification Tree

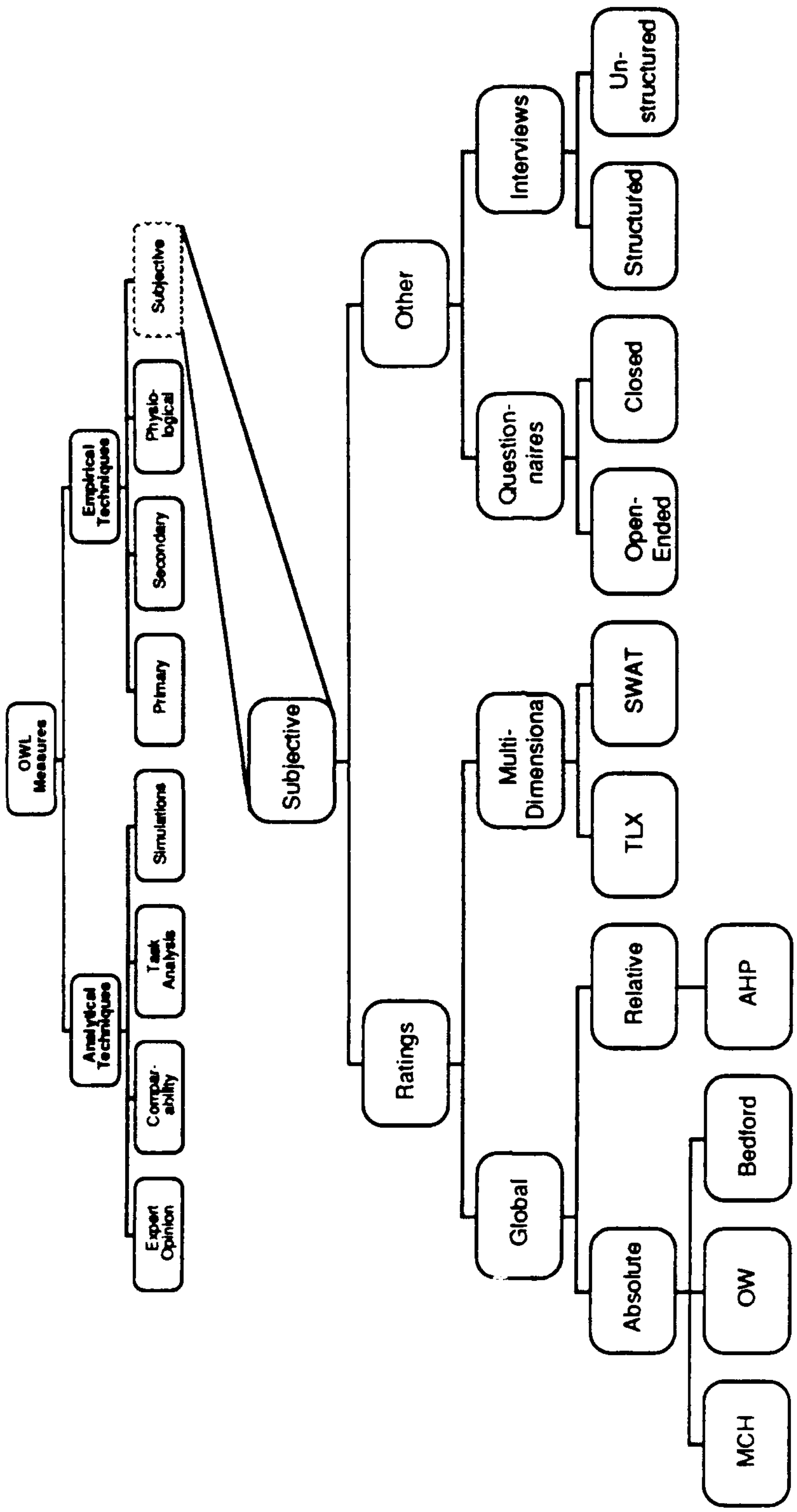


Figure 4-7. OWL Subjective Techniques Classification Tree

Knowledge rules

The expert system applies rules to determine the selection and presentation order of questions posed to the user and also to determine the selection and numerical ratings of the recommended techniques. These rules are normally hidden from the user (as are the thought processes of an expert); however, the user may opt to display them. The rules are specified as statements in the form: "If this premise is true, then perform this action or make this conclusion". Each rule is evaluated and when the current condition matches the premise state in the IF rule (i.e., the condition is TRUE), then the indicated action is performed.

The OWLKNEST rules are organized into rule groups corresponding to the decision tree classifications. OWLKNEST employs the strategy of initially assigning to all techniques the highest confidence value (equal 8), indicating that all techniques are potentially applicable to the user's assessment problem. Based upon the user's responses, the initial rule group prunes the classification tree by determining which branches of the tree, if any, can be eliminated (i.e., have their confidence values set equal to zero). A second group of rules refines the applicability rankings of the remaining techniques by setting confidence values to 2 for low applicability or to 5 for average applicability. Questions of resource availability primarily drive the elimination rules while the goals of the study drive the refinement rules. Only those questions that are needed to exercise the relevant rules will be presented to the user. For each unique workload study, there will be a unique set of questions posed to the user and a customized list of recommended techniques.

OWLKNEST Expert System Shell

The selection of an expert system shell for OWLKNEST was driven initially by requirements and constraints specified in the ARI OWL

program. Specifically, the selected expert system shell had to be a commercially available package with a cost less then $2,500. The software package had to come with unlimited or minimal cost distribution rights for run-only versions of the code. In addition, the expert system shell had to be compatible with the hardware and operating systems "typically" available to the Army user community.

Furthermore, the expert system shell selected had to accommodate the salient characteristics required by OWLKNEST. The knowledge representation scheme of the shell had to support the hierarchical structure of workload assessment techniques. The OWLKNEST domain is representative of a classical classification program, i.e, it attempts to characterize the user's problem and classify it in such a manner as to identify relevant recommendations. Classification problems are most amenable to data driven solutions that employ both forward and backward chaining inference capabilities. The desired output from OWLKNEST had to allow the ordering of recommendations into distinct categories of applicability and also provide multiple recommendations within a category. Also, OWLKNEST had to provide the user with a capability to do "what-if" types of analysis and to determine the impact of changing previous input data.

The characteristics of the approximately 25 expert system shells available for DOS systems were reviewed at a high level to eliminate those clearly inappropriate for OWLKNEST, based upon the above criteria. That screening process reduced the number of viable expert system shells to five. Using additional technical information provided by their vendors, these five software programs were analyzed in greater detail using criteria that addressed characteristics of the inference engine and of the user and developer interface. Finally, demonstration copies of each shell were obtained and used to create a prototype version of OWLKNEST.

Based upon the results of the detailed analysis and the experience of creating and using prototype versions of the OWLKNEST, one expert system shell clearly was superior to the others. That shell, called Exsys

Professional, was selected for OWLKNEST. While Exsys Professional, subsequently called the OWLKNEST shell or EXSYSP, was preferred over its competitors, it was not perfect for our purposes in all respects. In selecting this shell, we had to accept certain features that were not to our liking and which were, in some instances, not reliable. As appropriate and necessary, these limitations and constraints in the OWLKNEST shell will be described in subsequent sections of this handbook.

OWLKNEST Output

The output from OWLKNEST is a ordered list of appropriate OWL assessment techniques, each with a numerical rating value that reflects its applicability for the user's particular workload study. The final value of the numerical rating assigned to a given technique is based on the cumulative confidence or probability generated by the rules underlying each question and the responses selected by the user. The origin of the initial probabilities is the consensus of opinion formulated by the OWL Program team of workload experts.

These rating values serve as a guide to indicate the order in which the user should consider applying the technique. The user can optionally access the rules to see what parameters were influential in the determination of the listed ratings and the rating values assignments. Depending upon the user's responses to questions, it is possible that no recommendations can be made. In this case, OWLKNEST will briefly describe the situation and suggest alternative strategies for continuing, such as gathering additional details about the certain aspects of the workload problem

In using OWLKNEST, it is incumbent on the user to carefully consider which, if any, of the workload assessment techniques to implement from the list of recommendations. *OWLKNEST is a decision-aiding tool, not a substitute or replacement for the sound judgement of an analyst.*

Technical information sheets

To further assist the user in deciding which of the recommended techniques would be best suited, a one-page Technical Information Sheet (TIS) is provided for each technique. The TIS contains brief descriptions of the recommended technique(s) including implementation requirements, usage parameters, resource requirements, references, and points-of-contact (see Appendix A). Copies of the TISs can also be accessed from the operating system of the computer after terminating an OWLKNEST session within EXSYSP (see Part Six of this handbook).

Analysis of OWLKNEST output

OWLKNEST can be used in several different ways to provide insight on the appropriateness of various OWL assessment techniques at different stages in the development of a system. For example, OWLKNEST may be used to select workload prediction techniques during early system design efforts. Then after initial development is complete and a prototype is available, OWLKNEST might be used again to suggest workload techniques based upon currently available information and resources. Hence OWLKNEST can be used throughout the development cycle of systems.

OWLKNEST can also be used in a sensitivity analysis mode by changing one or more of the responses given. For example, in the first run, the analyst may choose to respond that no special equipment is available and obtain results based on that answer. In the next run, however, the analyst may want to see what other techniques would be recommended if audio and video recording equipment were available. In this case, the suggested list might include different techniques. For ease of comparability, OWLKNEST can generate a display of the previously recommended techniques along side the current results, each with their respective rankings. In this way, the analyst will be able to make informed decisions as to whether additional resources should be allocated to or required for the

workload assessment effort. This process is described in more detail in Part Seven of this handbook.

WC FIELDE: A Predecessor to OWLKNEST

OWLKNEST builds upon the foundation of a prior workload assessment tool, the Workload Consultant for Field Evaluation (WC FIELDE) (Casper, Shively, & Hart. 1986). WC FIELDE, developed by NASA, also utilizes an expert systems approach. It includes a number of rules which are used to rank 24 workload measurement techniques in terms of their appropriateness for the particular circumstances of the proposed study (Casper, Shively, & Hart, 1987). On the surface, OWLKNEST differs from WC FIELDE in two major ways.

1. The OWLKNEST knowledge base contains both analytical and empirical workload assessment techniques; WC FIELDE contains only empirical techniques.
2. OWLKNEST emphasizes those techniques suitable for operational and field testing, especially during the evaluation of Army systems; WC FIELDE does not.

In addition, there is a less visible but more fundamental difference between WC FIELDE and OWLKNEST. This difference reflects the strategy by which a final set of recommended techniques is determined. WC FIELDE begins with a blank slate, and through a dialogue with the user builds a set of recommended OWL assessment techniques; the user's answer to each question in a series of questions either increases or does not increase the likelihood that a particular assessment technique will be recommended.

The reverse approach is incorporated into OWLKNEST. With OWLKNEST, all assessment techniques that have demonstrated utility are initially included in the set of potential recommendations. Through a dialogue with the user, certain techniques are eliminated from future consideration. Then, additional dialogue with the user is used to refine the value to the user of the techniques which have not been eliminated. We

experimented with both of these approaches and found that the strategy built into OWLKNEST leads more quickly and efficiently to a final set of recommendations. We believe this difference is due principally to the fact that the OWLKNEST strategy presents fewer and less redundant questions to the user than does the WC FIELDE strategy.

PART FIVE: Installing OWLKNEST

Before You Begin

This chapter should be read carefully before using OWLKNEST. It provides information on the computer hardware and software that is needed to run OWLKNEST as well as instructions on how to install OWLKNEST on your computer.

Naming Conventions

CAPITAL letters are used within this text for command descriptions. Although either upper or lower case characters may be typed, any text shown in CAPITAL letters must be entered exactly as shown in the manual. Text shown in italics indicates something that the user will supply, e.g., a *filename*.

Hardware Requirements

To use OWLKNEST, an IBM PC or 100 percent compatible is required. The following minimum hardware configuration is recommended:

- An 80286 microprocessor-based computer equipped with a minimum of 640 Kilobytes (Kb) of random access memory (RAM);
- At least one floppy diskette drive of at least 360 Kb; and
- An additional floppy diskette drive or hard disk.

Software Requirements

PC or MS DOS 2.0 or a higher version of the operating system is recommended for OWLKNEST. Version 2.06 of the EXSYS Professional expert system shell was utilized in the development of OWLKNEST. Since

the OWLKNEST system is distributed under a runtime license, no additional software, such as the expert system shell, is required.

The Enter Key

The Enter or Return key is usually labeled with a bent left arrow (↵) on the keyboard and is used to indicate the end of a line. DOS will not process anything that has been typed until the Enter key is pressed. All the examples in this section will remind the reader that the Enter key must be depressed at the end of each DOS command.

OWLKNEST Software Installation

The OWLKNEST software is available on a set of three 5 1/4" low density (360 Kb) floppy diskettes.

OWLKNEST can be run from either floppy diskettes or installed on a hard disk. However, it cannot be executed from a single low density (360 Kb) floppy drive. If not installed on a hard disk, two floppy drives are required.

NOTE:	No software installation is required if OWLKNEST will be run only from floppy disks.

Hard disk installation procedure

NOTE:	The OWLKNEST installation procedure will create a subdirectory called OWLKNEST on the hard disk.

If installed on a hard disk, the OWLKNEST system requires a minimum of 700 kilobytes of storage. Therefore, at least this amount of space must be free on the hard disk.

To install the OWLKNEST software on a computer equipped with a floppy diskette drive and a hard disk drive, perform the following steps:

1. Power up the computer.

NOTE:	This procedure assumes that you will be in the root directory on the C drive. If an AUTOEXEC.BAT file automatically executes on your system and it contains commands that change to a subdirectory, enter the command **CD ** ↵ to return to the root level.

2. Wait until the operating system prompt (usually a C>) is displayed.
3. Insert the diskette labeled 'Disk 1' into the source floppy diskette drive. The source floppy diskette drive is the one that you will be inserting diskettes into. The left and the top diskette drives are generally the source drive referred to as drive A.
4. Change the default drive to the source floppy diskette drive by entering the drive letter followed by a colon, e.g., **A:**↵.

NOTE:	Do not type a colon after the drive letters in the following installation command line.

5. Type **INSTALL *C A COLOR*** ↵ where:

   C is the usual designation for the hard disk,

   A is the usual designation for the floppy diskette drive, and

   COLOR indicates the color monitor.

If your computer uses a different designation for the hard or floppy disk drives, use those designations rather than the ones indicated here. If you do not have a color monitor, type NOCOLOR rather than COLOR.

NOTE:	The installation procedures will cause a prompt to be displayed which indicates when the diskettes labelled 'Disk 2' and 'Disk 3' are to be inserted in the diskette drive. Make sure to insert the disks in the proper sequence. The installation procedure halts until any key on the keyboard is depressed to indicate that the proper diskette has been inserted in the drive and installation is to continue.

6. Follow the instructions displayed on the screen and insert each disk as requested.

DOS considerations

The CONFIG.SYS file should be reviewed to ensure that the following parameters are set to at least the values shown below:

FILES=25

BUFFER=20

To modify the CONFIG.SYS file, any text-oriented word processing or editor system can be used.

PART SIX: Operating OWLKNEST

This part of HOOT describes how to operate the OWLKNEST software including the various options available and their impact during the use of OWLKNEST. It also presents a checklist of information that may be requested by OWLKNEST along with a descriptive listing of the criteria corresponding to the checklist categories. Additionally, it explains the components of the OWLKNEST user-computer interface and how they are used. Finally, each OWLKNEST feature is described, step-by-step.

CAPITAL letters are used within this text for command descriptions. Although either upper or lower case characters may be typed, any text shown in CAPITAL letters must be entered exactly as shown in the manual. Text shown in italics indicates something that the user will supply, e.g., a filename.

For all of the descriptions and discussions of OWLKNEST presented in this handbook, it is important to distinguish those features of the expert system that are determined by capabilities and constraints of the expert system shell and those that were determined by the OWL program team which developed and encoded the workload-specific knowledge base into the expert system shell. The acronym EXSYSP is used to refer to features specifically determined by the expert system shell, EXSYS Professional. The acronym OWLKNEST refers to the OWLKNEST knowledge base and to the total expert system package that incorporates this knowledge base.

The OWLKNEST Checklist

Table 6-1 presents a checklist of the information and associated input parameters that can be used to define the desired characteristics of a particular workload study before starting OWLKNEST. The items on the checklist correspond to the criteria defined to identify key features of

Table 6-1. The OWLKNEST Checklist

Techniques:	❑ Analytical ❑ Empirical ❑ Both
Operator:	❑ Available ❑ Not Available
System/Hardware:	❑ Available ❑ Not Available
Subject Matter Experts:	❑ Available ❑ Not Available
Comparable System:	❑ Available ❑ Not Available
Time Constraints:	❑ None (or one of the following:)
Preparation:	❑ < 1 day ❑ < 1 week ❑ < 1 month ❑ > 1 month
Data collection:	❑ < 1/2 hour ❑ < 1 hour ❑ < 4 hours ❑ > 4 hours
Operator:	❑ < 1 minute ❑ < 5 minutes ❑ < 15 min ❑ > 15 minutes
Scoring:	❑ < 1 minute ❑ < 5 minutes ❑ < 15 min ❑ > 15 minutes
Data Analysis:	❑ < 1 day ❑ < 1 week ❑ < 1 month ❑ > 1 month
Ease of Use	❑ Preparation ❑ Collection ❑ Scoring ❑ Analysis ❑ Not an issue
Performance:	❑ System ❑ Operator ❑ Both
Spare Capacity Analysis:	❑ Yes ❑ No
Available Task Data:	❑ Descriptions ❑ General task times ❑ Estimated time of detailed tasks
Available Equipment	❑ Audio Tape ❑ Video Tape ❑ Pupil Diameter ❑ EKG ❑ EEG ❑ Oculometer ❑ IBM PC ❑ None
Workload Dimensions:	❑ Auditory ❑ Cognitive ❑ Motor ❑ Physical ❑ Stress ❑ Time ❑ Visual ❑ None
Operator Contact:	❑ Face-to-face ❑ Remote ❑ None
Real-Time Application:	❑ Yes ❑ No
Environment:	❑ Operational ❑ Laboratory ❑ Both
Training Time:	❑ < 15 mins ❑ < 1 hour ❑ < 4 hours ❑ > 4 hours
Operator Interference:	❑ None ❑ Minimal ❑ Not a concern
Operator Intrusiveness:	❑ None permitted ❑ Limited head ❑ Minimal ❑ Limited Eye ❑ Not a concern
Desired Outputs:	❑ Quantitative ❑ Qualitative ❑ Either
Diagnosticity:	❑ Global ❑ Detailed ❑ Either
Result anonymity:	❑ Yes ❑ Not a concern
Sensitivity:	❑ Large ❑ Both subtle and large

Table 6-2. OWLKNEST Criteria

Techniques:	Workload techniques have been divided into two general categories: 1. Analytical, or predictive, techniques that may be applied early in system design without an "operator in the loop". 2. Empirical techniques requiring an operator in the loop using a simulator, prototype, or representative system. In order for empirical techniques to be utilized, both an operator and a representative system must be available for use during the workload study. If neither is available for the study, then only analytical techniques will be considered.
Operator:	A person who can operate the system or uses the system output during the workload evaluation is required to use empirical techniques. If none is available, then only analytical techniques will be considered.
System Hardware:	System availability refers to whether a system, mockup, or simulator can be used by operators to perform the necessary tasks during the workload study. If none is available, then only analytical techniques will be considered.
Subject Matter Experts:	Subject Matter Experts (SME) are individuals who have extensive knowledge of the tasks and functions of the system that is under study, a predecessor, or one which is functionally similar. They may be used as sources of expert opinion and workload information.
Comparable System:	A comparable or predecessor system is one which is functionally similar to the system under study.
Time Constraints:	Time constraints are any time limits or requirements which may impact the study (e.g., total time for the workload study, time for a single decision or data collection trial). If there are no externally imposed time constraints, estimate the amount of time you can or want to spend. The time constraints are divided into the following: **Preparation** — total time spent by the workload practitioner in preparing for the workload study. It does not include the time spent training or preparing the operator. **Data Collection** — time required for the workload analyst to utilize or apply the technique (e.g., administer questionnaires, collect raw data, etc.) for a single session. **Operator** — time required by the operator to complete an OWL measure (e.g., a questionnaire). **Scoring** — time required to transform the collected data into a usable form for analysis. This might include changing rating scale marks to numerical scores or determining performance success or failure based on known criteria. **Analysis** — time available for data analysis including analyst's time to consolidate data, run statistical analyses, graph, or interpret the results for a workload study.

Table 6-2. OWLKNEST Criteria (Continued)

Ease of Use:	Indicates whether the ease of use of the technique is to be considered for the following areas: **Preparation** — advance preparation necessary by the analyst for the study, e.g., to learn how to use a technique. It does not include training or preparing operators in OWL assessments. **Data Collection** — utilization (or implementation) of the technique (e.g., administering questionnaires, recording observations, etc.). **Scoring** — transformation of the data collected into useable form for analysis (e.g., assigning numerical scores to qualitative data, determining performance success/failure based on known criteria). **Analysis** — data consolidation, statistical treatments, graphing, and interpretation of the results.
Performance:	System/hardware performance relates to overall success or failure in achieving the objective. Operator performance relates to specific behaviors or tasks that the operator performs. While system performance measures may be less difficult to obtain (e.g., mission success), they may only provide information about work overload. Operator performance may be more difficult to measure, but may provide more information about varying levels of workload.
Spare Capacity Analysis:	The human may be viewed as having a limited capacity or ability with which to process information. A simplistic example is an operator, who currently using only 25% of capacity, has 75% spare capacity to apply to an additional task, increased task demands, emergency situations, etc. The concept of workload can be defined in terms of the relationship between resource supply and task demand. Changes in workload may result from fluctuations of operator capacity or task resource demands.
Available Task Data:	Workload techniques have various requirements for the level of specificity of task data including: **Task descriptions** — general descriptions of what tasks the operator will be performing, **General task times** — gross estimates of time to accomplish general level tasks, and **Detailed task times** — time estimates of specific tasks, to include the order in which the tasks should occur.
Available Equipment:	Indicates the special equipment that the user either has access to or will be able to acquire.

Table 6-2. OWLKNEST Criteria (Continued)

Workload Dimensions:	Workload techniques vary in appropriateness for various dimensions such as: **Auditory** — sensing (receiving) information from auditory sources (e.g., speech communication, signals, alarms). **Cognitive/mental** — planning, prediction, calculation, information absorption and processing. **Motor/psychomotor** — writing, tracking, activating control mechanisms (e.g., button pushing, keyset entry). **Physical** — gross motor activity such as manual handling or movement of materials. **Stress/frustration** — condition, circumstance, task or other factors with understandable physiological or psychological consequences to the individual. Frustration may be viewed as dissatisfaction arising from unresolved issues. **Time** — usually expressed as a ratio of time required for task or mission divided by time available. **Visual** — sensing (receiving) information from visual sources (e.g., visual display terminals, graphic or alphanumeric materials, warning lights).
Operator Contact:	Type of contact between the operator and the data collector during data collection. It does not include contact during any training.
Real-Time Applications:	Real-time application means that the technique is used practically simultaneous with the occurrence of the task or event to which it is applied. If not real-time, the application of the technique is delayed for some period of time after the task or event has occurred.
Environment:	Laboratory indicates a lab setting and situation. Operational indicates an actual field setting.
Training Time:	Indicates the amount of time available for the operator to be trained to perform the workload technique.
Operator Interference:	The degree to which the workload technique interferes with the performance of the operator's primary tasks.
Operator Intrusiveness:	Intrusive refers to the degree to which the application of the workload technique invades the human body. An example is sensors attached to the body for monitoring heart rate.
Desired Outputs:	Quantitative outputs are expressed in numerical terms; qualitative outputs are expressed in only verbal terms.
Diagnosticity:	The extent to which a technique reveals not only the overall assessment of OWL (global sources) but also information about component factors (detailed sources). For example, a technique that differentiates among various sensory, perceptual, cognitive, and psychomotor aspects of human performance is considered to reveal detailed sources of workload.
Anonymous Results:	Anonymous outputs (results) are those for which the source of the data (subjects/operators) is not identified or is kept confidential.
Sensitivity:	Sensitivity of workload techniques is the degree to which the various techniques can differentiate among levels of load placed on the operator. It also depends on the appropriateness of the technique for the system.

workload techniques as described in Part 2. Table 6-2 presents a brief description of each of the OWLKNEST criteria. The exact set of questions and the order in which they will be shown to the user varies since the questions selected to be displayed are limited to only those that will provide user-supplied data needed by OWLKNEST to apply to a rule. This approach attempts to quickly focus on the most applicable technique(s) and minimize the number of questions posed by the system.

How to Start OWLKNEST

OWLKNEST can be run from either a hard disk or from floppy diskettes as described below.

From hard disk

NOTE:	This procedure assumes that you will be in the root directory on the C drive. If an AUTOEXEC.BAT file that is automatically executed contains commands that change to a subdirectory, enter the command **CD** \ ↵ to return to the root level.

OWLKNEST can be run from a hard disk only if previously installed as described in Part 5. The procedure to start OWLKNEST from a hard disk is described below:

1. Enter **CD OWLKNEST↵**
2. After the DOS prompt, type **OWLKNEST↵**

These commands change control to the subdirectory named OWLKNEST where the batch file also named OWLKNEST can be executed.

From dual floppy (360Kb) diskettes

The procedure to start OWLKNEST from dual floppy diskettes is described below:

NOTE:	Disk 3 contains the OWLKNEST Technical Information Sheets which are not required to run OWLKNEST.

1. Insert disk 1 into drive A and disk 2 into drive B.
2. Make drive A the default drive by entering **A:↵**

NOTE:	A different batch file named OWLKNES1, rather than OWLKNEST, is used when running from duel floppy diskettes to distinguish the software on drive a and b.

3. Type **OWLKNES1↵**.

EXSYSP User-Computer Interface

EXSYSP incorporates two user-computer interface features with respect to the use of the Enter or Return key (sometimes labeled as ↵ on the keyboard). One requires the user to respond to a question without using the Enter key while the other does require the use of the Enter key.

Do not use the Enter key

The Enter key is not required when the user is selecting a single character response from a displayed list of fixed options. For example, a question is posed to the user and the answer can either be yes or no (usually displayed as (Y/N)). In this case, the user must only enter the character corresponding to the desired option **without** using the enter (↵) key. Depression of only the Enter (↵) key will result in the default option being selected.

Inappropriate use of the Enter key

If the Enter (↵) key is depressed in addition to the appropriate character, the Enter (↵) key will be entered into the computer's input buffer and the default option will be selected for the next question without the question ever being displayed for the user.

Must use the Enter key

The Enter key must be depressed when variable inputs are required or permitted. For example, entering a 1-8 character filename or selecting either a response to a question or a command option. In this case, the user must terminate the input line by the depressing the Enter key (↵).

The commands in this section that require the use of the Enter key will be labeled by showing the ↵ symbol. Where the Enter key is not required, a brief note about the input options will be displayed.

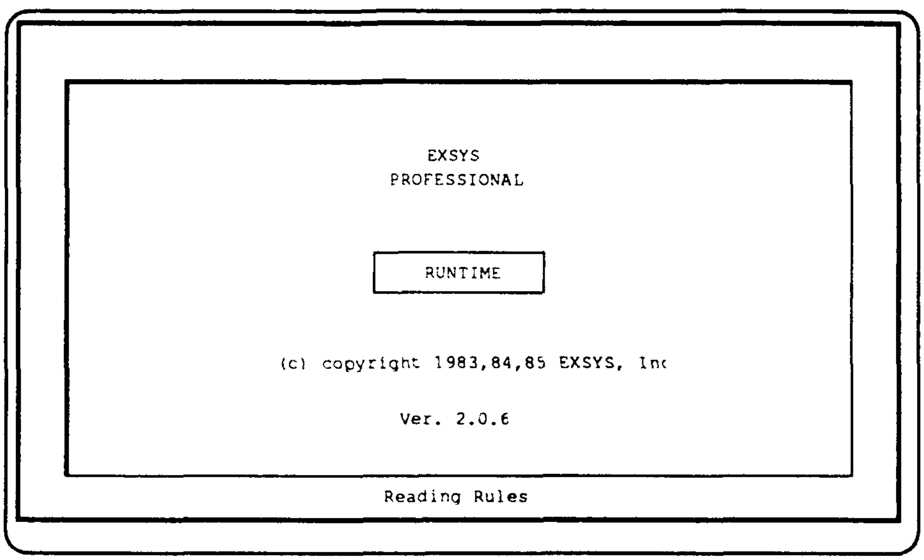
Initial Screens

A series of screens, as described below, introduce OWLKNEST and allow the user to obtain detailed instructions on using the expert system software. These screens will be displayed each time OWLKNEST is run.

Introductory screen

While OWLKNEST is loading, the opening screen shown below will be displayed indicating the operating version of EXSYSP. The message 'Reading Rules' will be displayed in the middle of the bottom line on the screen while OWLKNEST is loading the rules. The loading time is dependent upon the speed of the processor in your particular computer. If the Enter key is depressed during this time, the default options will be automatically selected for the next response.

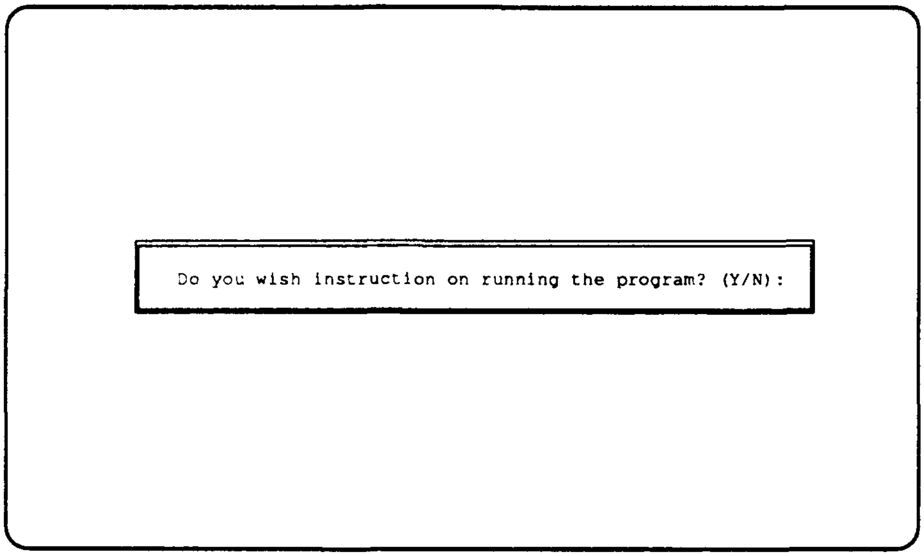
When finished reading rules, the EXSYSP Instruction screen will be displayed as described below.



EXSYSP instruction screen

NOTE:	Do **not** follow the entry of **Y** or **N** with the enter (↵) key.

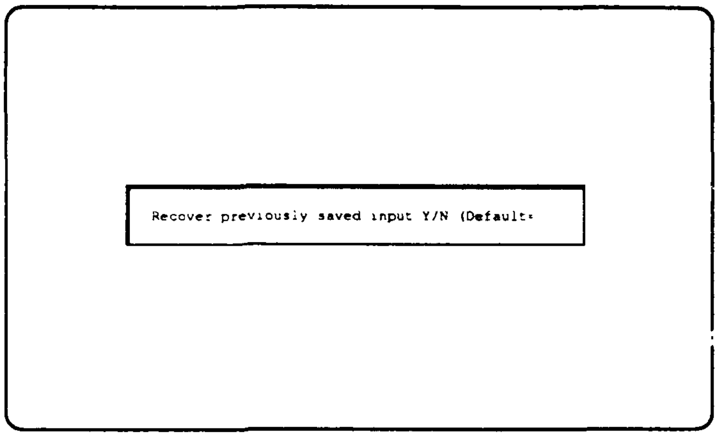
The EXSYSP instruction screen allows the user to obtain instructions on using EXSYSP. If the user types a **Y**, general information on using the EXSYSP expert system software is displayed. Any other entry (including the enter (↵) key) bypasses this option.



As will become evident to the user who exercises this option, the instructions provided were prepared by the vendor for EXSYSP and not the developers of OWLKNEST. Therefore, the instructions refer to using all the features available in the EXSYSP software, and they are not necessarily relevant to the OWLKNEST application of this expert system shell.

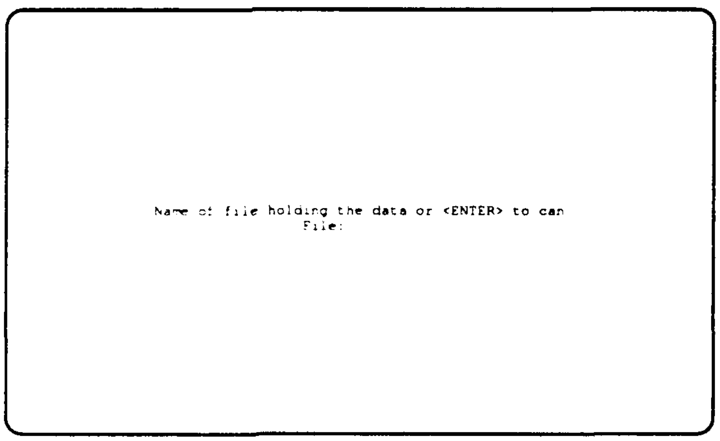
Recover data screen

The Recover Data screen allows the user to access data files saved through use of the Quit command from a previous OWLKNEST session (see the section below on quitting OWLKNEST and saving data). This option allows the user to continue processing a problem at the same point where processing was previously stopped.



NOTE: Do **not** follow the entry of **Y** or **N** with the enter (↵) key.

If the user types a **Y** to indicate that the saved data is to be recovered, then the user will be requested to enter the filename containing the saved data. (Valid EXSYSP filenames are described below in the section on managing OWLKNEST user files.) Any other entry (including the enter (↵) key) bypasses this option. After reading the saved data, OWLKNEST will return to the point where the Quit command was entered.



If the user incorrectly enters a filename that does not identify a previously saved file of input data, OWLKNEST will so inform the user. The user is then given the option to enter a different filename or to cancel the recovery process. The procedure for viewing the names of previously

saved input data files is described below in the section on managing OWLKNEST user files.

Title and author screen

The title and author screen displays the name of the system and the author. The depression of any key clears this screen.

```
EXSYS Pro

        Operator Workload Knowledge-based
        Expert System Tool (OWLKNEST)

        Army Research Institute Field Unit
                Fort Bliss, Texas
                       and
                 Analytics, Inc.
           Willow Grove, Pennsylvania

              Press any key to start:
```

Subject screen

The subject screen presents a brief description of the expert system, its version and release date, and an indication of who to contact for additional assistance. The depression of any key clears this screen.

```
Operator Workload Knowledge-based Expert System Tool (OWLKNEST)

Provides guidance in selecting the most appropriate technique(s)
to use for predicting and evaluating Operator Workload (OWL).
For additional information, consult the Handbook for
Operating the OWLKNEST Technology (HOOT), a user's guide for
OWLKNEST.
---------------------------------------------------
Version 1   (February, 1991)
---------------------------------------------------
If you have any questions or comments,
please contact Richard E. Christ (915) 568-4491


                    Press any key to start
```

At this point, OWLKNEST will either begin the question and answer dialogue to obtain necessary inputs from the user or, in the case of file recovery, return to the point where the user entered the Quit command.

User-Computer Dialogue

After the series of initial screens displayed whenever OWLKNEST is started, the system gathers information from the user about characteristics of the workload problem. This information is used in conjunction with the OWLKNEST knowledge base and the EXSYSP inference engine to determine which OWL assessment techniques to recommend. During the process of gathering information and determining recommendations, OWLKNEST presents or makes available to the user six distinctively different types of dialogue screen, each with their respective menu of options. From the user's perspective, these dialogue screens permit the user to engage in the following types of activities:

1. Answering questions,
2. Getting help,
3. Interpreting results and recommendations,
4. Examining the impact of alternative data,

5. Getting explanations, and
6. Saving data and quitting OWLKNEST.

The next six sections of this part of HOOT will describe and briefly discuss how to use the dialogue screens to engage in each of these user activities.

Answering Questions

OWLKNEST presents a series of multiple choice questions to the user. The specific questions and their sequence of presentation will vary depending upon the answers given to previous questions. A sample question and answer screen is illustrated below.

```
EXSYS Pro          You may select ONLY ONE value
        Workload assessment techniques to be considered in the
        analysis are
           1   Analytical techniques only
           2   Empirical techniques only
           3   Both analytical and empirical techniques
           4   I don't know
           5   Not applicable









                          ↕▶▶ :
Enter the value number(s) or select with arrow keys and press <ENTER>
WHY-rule used  <?>-details  QUIT-save  <H>-help
```

NOTE:	Here and elsewhere, while certain features of a dialogue screen are specified by the developers of OWLKNEST, others are dictated by the expert system shell, i.e, EXSYSP. This division of control over wording and format may lead to problems which can best be avoided by following the guidance contained in this manual (which is controlled by the developers of OWLKNEST) rather than the guidance suggested by the user-EXSYSP interface.

The user selects one or more of the numbered answers in response to each question. While the menu at the bottom of all question and answer

screens indicate that the user can select alternative answers either by typing the number(s) of the response(s) or by using the up and down arrow keys to highlight those numbers, there are several problems associated with using the arrow keys. Consequently, it is suggested that the user enter the number(s) of the selected answers(s) by typing these value(s) with the computer key pad and that the arrow keys not be used for this purpose.

The number of answers allowed for each question is shown at the top of the question and answer screen. To enter multiple answers, type the numbers associated with the desired answers, separating successive answers with commas or spaces. In the screen illustrated above, only one answer is permitted. The typed entry or entries (in the case of multiple answers) are displayed in the bottom middle of the screen, following the double arrowheads. Once the number(s) of the selected answer(s) have been typed and checked for their accuracy, they are entered by pressing the enter (↵) key. In the illustrated example screen, a 1 has been entered to indicate that only analytical techniques are to be considered for the workload problem under analysis.

I don't know and not applicable responses

The last two responses to each question can be used to indicate that none of the options are applicable to the problem at hand. Selection of the numbered response for the "I don't know" option is used when the user does not possess sufficient information at the present time to answer the question. The "Not applicable" choice is used to indicate that the question is not pertinent to the problem. These last two choices are used to prevent the expert system from forcing the user to select a clearly inappropriate alternative and subsequently using an inappropriate response in ranking the OWL techniques. While these two response alternative are treated the same way in OWLKNEST, the user's selection of one or the other response is stored in the data file to indicate quite different states of user knowledge of the workload problem.

Correcting typing errors

If a typing mistake is made and detected prior to depression of the enter key, it can be corrected by using the backspace key to delete the erroneous response and then typing the correct response.

Handling entry error

If the expert system detects an input error, an appropriate error message may be displayed at the bottom of the screen. Input errors include the following:

- A non-numeric key is entered,
- No input supplied, i.e., only the enter (↵) key is depressed,
- The entered number is outside the range of the displayed list, or
- Multiple inputs are entered when only one is permitted.

Even when an error message is not displayed, as will be the case following the first three input errors listed above, the error will cause the question and answer screen to be redisplayed, prompting the user to reenter the response.

Inconsistent responses

An inconsistent response is one that generally makes no sense. It involves making contradictory statements in response to a question permitting multiple answers. An example would be to enter 1 and 4 in response to the question shown below which indicates that available task data includes both 'Task descriptions' and 'No task information'.

```
EXSYS Pro          You may select any number of values

        Aavilable task data includes
          1   Task descriptions
          2   General task times
          3   Estimated time of detailed tasks
          4   No task information
          5   I don't know
          6   Not applicable

                    ↕ ▶▶ 1,4
Enter the value number(s) or select with arrow keys and press <ENTER>
WHY-rule used  <?>-details  Quit-save  <H>-help
```

Exsysp does not include any provisions for detecting an inconsistent response. Therefore, the user is cautioned that entry of any inconsistent responses may result in inappropriate firing of rules and provide contradictory or confusing results. In the event that such results are obtained, it is recommended that all inputs be reviewed to determine if any inconsistencies were specified.

Command options for question and answer screens

At the bottom of the question and answer screen, additional commands are available to the user as summarized below:

- WHY — provides explanation by displaying the rule(s) associated with a particular question,
- ? — provides OWLKNEST developer help for interpreting OWL questions and answers,
- Quit — allows the data entered to be saved and optionally ends the OWLKNEST session, and
- H — provides EXSYSP help for interpreting commands and logic in the OWLKNEST shell.

NOTE:	Do not follow entry of either single character response, i.e., ? or H, with the enter (↵) key response.

The usage and capabilities of these commands are described in subsequent sections. These commands can be typed in either lower or upper case.

Getting Help Information

Two types of help information are available:

1. Entry of ? displays OWLKNEST-specific help information and
2. Entry of **H** displays EXSYSP-specific help information.

NOTE:	Do not follow entry of either **?** or **H** with the enter (↵) key.

OWLKNEST-specific Help

The OWLKNEST-specific Help files are available for each question. These files were prepared by the developers of OWLKNEST. They present additional details to clarify the meaning of questions and answers, and may also describe how the selection of certain alternatives affect the results. To view the help file associated with an OWLKNEST question, the user types a ? at the prompt. The sample below illustrates how to get detailed information about the operator availability question.

```
EXSYS Pro        You may select ONLY ONE value

        Operator availability to interact with equipment for t
        workload study is
          1   Operator(s) available
          2   None available
          3   I don't know
          4   Not applicable

                              ↕ ▶▶ ?
Enter the value number(s) or select with arrow keys and press <ENTER>
```

A sample OWLKNEST Help screen is shown below:

```
A human subject who can operate the system during the
workload evaluation is required to use
empirical techniques.  If none is available,
then only analytical techniques will be considered.











Return-<ESC> :
Press <ENTER> to continue:
```

After viewing the displayed detailed information, the user can return to OWLKNEST by hitting **<ESC>**. Subsequent pages of the help screens will be displayed if the enter (↵) key is depressed. After the help screen displays are completed, the original question screen where the user requested OWLKNEST help will be redisplayed.

EXSYSP-specific help

The EXSYSP Help screens are available to explain EXSYSP commands and features. The information presented was prepared by the developers of EXSYSP, and is the same for each question and answer screen. To view this information the user types an H as a response to a question screen. After viewing the information the user can return to OWLKNEST by hitting **<ESC>** at any time or display subsequent pages of the help screens with the enter (↵) key. A sample EXSYSP Help screen is presented below.

```
The computer is asking you for input to give it the data it
needs to determine which of the posssible answers is most
appropriate.  You will be presented with a phrase followed by
a numbered list of possible completions of the phrase.  Select
the item(s) from the list that are appropriate for your problem
and input the numbers.  If more than one item is appropriate,
separate the numbers witha a comma or space.

Return-<ESC> :
Press <ENTER> to continue:
```

Some EXSYSP Help screens contain text with highlighted key words. The highlighted key words can be linked to other screens in the EXSYSP help file which contain additional relevant information. The user activates this feature by typing <F1> (the function key labeled F1), causing the highlighted key word in the text to be displayed at the bottom of the screen. Repeated depression of <F1> causes the display to scroll through subsequent key words. When the key word for which additional information is desired appears at the bottom of the screen, depressing the enter key displays the associated information. The ESC key is used to return to the original question screen.

If you do not understand why the computer is asking you this question, you can ask it what rule it is trying to apply by typing "**WHY**" and then pressing the <ENTER> key. The computer will respond by displaying the **rule** it is trying to determine the validity of.

You will notice that the **qualifier** it was asking you for is in the rule's **IF part**. Press <ENTER> and the computer will either re-ask you the original question or display another **rule**. If another rule is displayed, it is because the first rule shown was only being used to derive information for another rule. The computer will continue displaying rules until it reaches the base rule it was trying to apply. This rule will have one or more choices (the possible solutions to the problem).

Keyword Information-<F1> Prev. Screen-<Page UP> Return-<ESC> :
Press <ENTER> to continue:

Interpreting Recommendations and Results

After the OWLKNEST user dialogue is completed for a particular workload study, OWLKNEST displays its recommendations, as shown in the example screen below. The EXSYSP term "choice", as presented on the options list at the bottom of the screen, refers to the recommended workload categories and techniques.

```
EXSYS Pro                        ** RESULTS **
                                                          VALUE
   1 Analytical                                              8
   2 Expert Opinions                                         8
   3 Interviews                                              6
   4 Questionnaires                                          6
   5 *Closed Questionnaire                                   6
   6 *Delphi Interviews                                      5
   7 Prospective Ratings                                     5
   8 Pro-Global                                              5
   9 *Pro-OW                                                 4

 All choices <A> only if value>1 <G> Print <P> Change and rerun
 Rules used <line #> Quit/save <Q> Help <H> Done <D>:
```

The recommended techniques are listed in the order of their value for the particular workload study. The most highly recommended techniques are listed first followed by those of lower value. The numeric values shown for each technique have the following general interpretations:

7-8 High recommendation

4-6 Average recommendation

2-3 Low recommendation

The higher the rating values, the more appropriate the technique is to the specific assessment situation. These rating values are based on probabilities built into the expert system rules. Since the ratings are based upon informed judgment, they serve as a guide to indicate the order in which the user should consider applying the technique. It is incumbent upon the analyst using OWLKNEST to use judgment in choosing which of the listed techniques to assess workload.

The format and structure of the list of recommendations is dictated by EXSYSP. It does, however, include all nodes of the workload assessment hierarchy. Those representing the lowest node of the tree, typically corresponding to specific assessment techniques, are preceded by asterisks (*). Higher-level nodes that represent categories are shown because OWLKNEST may be able to decide that while a particular class of

techniques appears most suitable, that class does not contain a specific technique that meets all the needs and objectives. For example, OWLKNEST may have determined that task analysis techniques would best meet the needs of a particular situation, but all of the specific task analysis techniques contained in the OWLKNEST knowledge base may require more information or resources than the user indicated were available. In this case, the user at least has an indication of the kind of technique that should be further pursued.

Command options for results screens

NOTE:	Do not follow entry of any single character option with an enter ((↵)) key response. Only entry of the <line #> option requires use of the enter (↵) key.

The command options displayed at the bottom of the results screen perform the following functions:

A	Displays all techniques that are elements of rules activated by the inference engine, including those techniques subsequently eliminated and having a value of zero.
G	Displays only those techniques which are recommended for consideration, and hence having a value greater than zero.
P	Prints the results with optional storage of results in a disk file. Described in detail below in the 'printing results' subsection.
C	Changes one or more of the input values and reruns OWLKNEST. Described in detail in the section below in the 'impact of alternative data' section.
line #	Displays all rules associated with the technique shown on the indicated line number of the results screen. Described in detail in the section below on getting explanation.
Q	Optionally saves input to a file, exits OWLKNEST, or both. Described in detail in the section below on quitting OWLKNEST and saving data.
H	Displays EXSYSP help information for various commands. Described above in the section on getting help.
D	Indicates current analysis is completed. When selected from the command menu on the results screen, the OWLKNEST

session is ended without any opportunity to save the input data. See the section below on getting explanation for another use of this command option.

Printing results

The selection of the P (Print) command from the results screen menu allows the user to print a copy of the results or to store the results in a file. The dialogue screen shown below is displayed to the user after the Print command is selected.

```
Do you wish to also print the data you input? (Y/N) (Default=Y)

Print to printer <P> or disk file <D> (Default=P):


Turn printer on.  Press any key when ready:
```

As illustrated, the user is presented with the following options:

1. Print the user's input data as well as the results, and
2. Select printer or disk file output.

If the disk file output option is selected, the user will be asked to enter a filename. This filename should be distinctively different from the filenames created when only the input data are saved in conjunction with the quitting OWLKNEST option (described below). Valid EXSYSP filenames are described in the section below on managing OWLKNEST user files.

If the printer output option is selected, a message will be displayed indicating that the printer should be turned on. When the printer is ready, press any key to obtain a printout.

NOTE:	To execute the printer option the user must have an appropriate printer correctly linked to the OWLKNEST system through the output port of the personal computer.

Examining the Impact of Alternative Data

The C (Change and rerun) option of the results screen allows the user to modify one or more of the input values and rerun OWLKNEST in order to determine the impact on the recommended list of techniques. The procedures for utilizing this capability are described in this section while strategies for effectively employing this capability are presented in Part Seven.

Upon selecting the C command, the user is presented with a new menu option at the bottom of the results screen, as illustrated below. This option allows the user to indicate whether the current values should be stored in order to compare them with the new results. It is recommended that the user select the default value.

```
EXSYS Pro                    ** RESULTS **
                                                            VALUE
  1 Analytical                                                8
  2 Expert Opinions                                           8
  3 Interviews                                                6
  4 Questionnaires                                            6
  5 *Closed Questionnaires                                    6
  6 *Delphi Interviews                                        5
  7 Prospective Ratings                                       5
  8 Pro-Global                                                5
  9 *Pro-OW                                                   4

Do you wish to store the current results for comparison with the new
results you will be calculating? (Y/N) (Default=Y) :
```

Next, screens labeled 'CHANGE INPUT DATA' will be displayed that show the current value for each input item. The display shows a sequential list of all input items along with their current values (the list may

not correspond to the order in which the questions were originally posed). The following four command options are presented in a menu at the bottom of this screen:

- Enter number of line to change;
- <O>, to delete changes and restore original data;
- <R>, to rerun OWLKNEST with new data; and
- <H>, for EXSYSP help (see section above on getting help).

To change any input value, enter the number that appears next to the input data item followed by the enter (↵) key. The original question and answer screen that corresponds to the item will be redisplayed and new response value(s) can be selected using the procedures described in the 'Answering Questions' section. Modification of the input values may generate additional questions that were not originally displayed when this particular workload problem was previously analyzed. After all necessary input values have been supplied for the changed input, the 'CHANGE INPUT DATA' screen will be redisplayed and the user can modify other input values as appropriate. After all desired changes have been completed, enter the **R** option to rerun the modified data. Note that while the modified data is processed, EXSYSP causes the monitor screen to go blank. The modified results are displayed as described above except that now, if the old values were saved for comparison, two columns of results are shown, one listing the original recommendations and values, and the other listing the results produced by the modified input data.

Repeated changes and reruns

The Change and rerun option of the results screen can be repeated as often as desired. However, after the first use of the option, the user is given an additional command menu that introduces the following three choices to specify which, if any, earlier results will be compared with the results about to be produced by the newly modified input data:

N Store and display the most recent previously generated results to compare with the new results.

R Store and display the original results to compare with the new results, and

C Do not store or display any results to compare with the new results.

Whenever initiating the change and rerun option, entering O will restore the original values to the list of all the original input items.

Getting Explanations

At practically any point while the user is answering questions posed by OWLKNEST or interpreting the guidance provided by the OWLKNEST results, the user may wish or need to understand the reasoning process used by OWLKNEST. In short, the user may want an explanation for why a particular type of data is needed or how a particular technique has come to be recommended.

On both the question and answer screens and the results screens, command options are provided which allow the user to see the rules that are currently under evaluation. The menu presented with the question and answer screens includes the command 'WHY', which, if selected, displays the rules associated with a particular question. The menu presented with the results screen includes the command 'Rule used <line #>', which, if selected, displays all rules associated with the technique shown on the indicated line number of the results screen.

The question screen shown below illustrates the WHY command available to the user.

```
EXSYS Pro            You may select ONLY ONE value
        Subject matter expert (SME)/experienced operator
        availability for use during the workload study is
          1   Subject matter experts/experienced operators available
          2   None available
          3   I don't know
          4   Not applicable

                         ↕▶▶ Why
Enter the value number(s) or select with arrow keys and press <ENTER>
WHY-rule used  <?>-details  Quit-save  <H>-help
```

The following rule displaying screen appears after the user types WHY and depresses the enter ((↵)) key.

```
EXSYS Pro            RULE NUMBER: 48* (alim nosme)
  IF:

       (1) Analytical- Conf. > 0/10
   AND (2) Subject matter expert (SME)/experienced operator availability for
           use during the workload study is None available

  THEN:
           Expert Opinions - Confidence=0/10
   AND     Interviews - Confidence=0/10
   AND     Prospective Ratings - Confidence=0/10
   AND     Pro-Global - Confidence=0/10
   AND     Pro-Multidimensional - Confidence=0/10
   AND     Questionnaires - Confidence=0/10

  NOTE: Rule out all techniques that require a subject matter expert if none
        is available.

 IF line # for derivation, <K>-known data, <C>-choices
↑ or ↓ - prev. or next rule, <J>-jump, <H>-help or <ENTER> to contine: 1
```

By examining the second premise of the rule show above, the user may determine the consequences of not having subject matter experts available for use during the workload study.

The exact same rule screen illustrated above (i.e., the screen displaying Rule Number 48) could also be presented in response to a need for explanation that arises while the user is attempting to interpret recommendations given in a results screen. In this case, the user would type in the line number of a results screen that corresponded to the listing of, say, the class of techniques called 'Expert Opinion', when that class of techniques is assigned a rating value of zero. (Note that the situation described here presumes that the user has also selected the command <A> to display all techniques considered by OWLKNEST for this workload

problem.) In this latter example, the user could determine that expert opinion is not a recommended technique for the workload study, at least in part, because no subject matter experts are available.

The examples given above illustrate how the user can get and interpret OWLKNEST-provided explanations for why questions are asked and how techniques are assigned specific values in the results. A complete description of all the rules and their possible interpretations is beyond the scope of this handbook. The current version of OWLKNEST incorporates over 200 rules and, depending upon the particular workload problem, many rules may be linked or chained together while the expertise encoded in the OWLKNEST software program evaluates user input data to produce its recommendations.

If a color monitor is used, EXSYSP color codes each If- premise of the rule as follows:

Yellow:	Indicates the condition is true,
Red:	Indicates the condition is false, and
Blue:	Indicates the interpretation of the rule is unknown at this point in time.

However, our experience has shown that the color actually associated with a rule premise may not be a reliable reflection of the input data. Consequently, the user is encouraged not to rely solely on this color coding scheme in interpreting the outcome of a rule evaluation.

Command options for rule screens

The command options available with the screen which displays rules associated with a given question or a specified recommendation are listed and briefly described below:

NOTE:	Do not follow entry of **K**,**C**, **J**, or **H** with the enter (↵) key.

line #	Displays the derivation of the data for the IF conditions as described below.
K	Displays all currently known data as described below.

C	Displays current values for all choices (or OWL techniques) with non-zero values as described below.
J	Jump to another rule.
Up arrow	Display the previous rule. The previous rule is the one sequentially before the current rule in the rule set. It is not necessarily the rule that was previously applied.
Down arrow	Display the next rule. The next rule is the one sequentially after the current rule in the rule set. It is not necessarily the rule that will be next applied.
H	Display EXSYSP help information as previously described in the 'Obtaining Help' section.
Enter (↵) key	Return to question and answer dialogue.

Derivation of data

The derivation of the data being used to evaluate the current rule can be obtained by typing the line number (indicated in parenthesis) of the item of interest followed by the enter (↵) key. For example, in the screen illustrated below, the first clause in the IF statement 'Analytical-Conf > 0/10' is line number 1. The results of entering <1> to do a data derivation query for this premise are shown at the bottom of the screen below. Typing any character will return the user to the previous screen.

```
EXSYS Pro              RULE NUMBER: 48  (alim noame)

 IF:

      (1) Analytical- Conf. > 0/10
  AND (2) Subject matter expert (SME)/experienced operator availability for
          use during the workload study is None available

 THEN:
          Expert Opinions - Confidence=0/10
  AND     Interviews - Confidence=0/10
  AND     Prospective Ratings - Confidence=0/10
  AND     Pro-Global - Confidence=0/10
  AND     Pro-Multidimensional - Confidence=0/10
  AND     Questionnaires - Confidence=0/10

 NOTE: Rule out all techniques that require a subject matter expert if none
       is available.

 The value of the choice is 8
 Press any key to continue:
```

Known data

The known data screen displays the current value of all user inputs entered up to the point of the the user request. It is obtained by typing a **K** while viewing the rule screen. The known data screen displays all the qualifier data currently known by OWLKNEST.

```
EXSYS Pro                     ** KNOWN DATA **

 1  Workload assessment techniques to be considered in the
    analysis are Analytical Techniques only
 2  Indicates if analtical techniques should be considered -
    yes
 3  Indicates if empirical techniques are appropiate = no
 4  insufficient information available to make any
    recommendations -  no

 All choices <A>  only if value>1 <G>  Print <P>
 Rules used <line #>  Quit/save <Q>  Help <H>  Done <D>:
```

Choice status

The Choice Status screen is generated by typing a **C** while viewing a rule screen. The choice status screen lists the choices still valid or under consideration as workload techniques used by OWLKNEST. EXSYSP choices correspond to the OWLKNEST workload techniques.

```
EXSYS Pro                       ** CHOICE STATUS **
                                                                  VALUE
1  Analytical                                                       8
2  Expert Opinions                                                  8
3  Interviews                                                       6
4  Questionnaires                                                   8
5  Prospective Ratings                                              8
6  Pro-Global                                                       8
7  Pro-Multidimensional                                             8
8  *Comparability Analysis                                          8
9  Task Analysis                                                    8
10 Task-Based Task Analysis                                         8
11 Time-Based Task Analysis                                         8
12 Simulations                                                      8

All choices <A>  only if value>1 <G>  Print <P>
Rules used <line #>  Quit/save <Q>  Help <H>  Done <D>:
```

Command options for the known data and choice status screens

The screens displayed in response to selecting the <K> known data and <C> choice status options available on the rule screens each, in turn, display a menu of command options. These later options are listed and briefly described below.

A Has no function for the K screen but for the C screen this option displays all techniques that that are elements of rules activated by the inference engine, including those techniques subsequently eliminated and having a value of zero.

G Has no function for the K screen but for the C screen this option displays only those techniques which are recommended for consideration, and hence having a value greater than zero.

P Prints the information displayed on the K or C screen. (See detailed description in the subsection above on printing results.)

line # For the K screen entering a number displays the derivation of the known data that has that line number in the screen. (See details above in the subsection on derivation of data.) For the C screen entering a line number displays all rules associated with the technique shown on the indicated line number of the screen. (See details in the section above on getting explanation.)

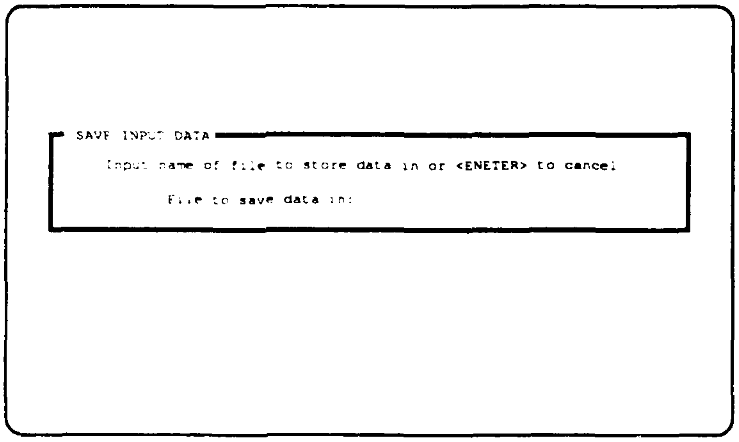
Q Optionally saves input up to this point to a file, exits the K or C screen and returns to the rule screen from which the K or C option was selected, or both.

H Displays EXSYSP help information for various key words.

D Indicates current K or C analysis is completed. When selected from either the K or C screen the rule screen from which the K or C option was selected is redisplayed. (Note that the only difference between the Q and D options on the K and C screens is that the Q option allows data and results to be saved while the D option does not.

Quitting OWLKNEST and Saving Data

The Quit or <Q> command allows the user to end the OWLKNEST session and optionally save the current data for later use. The screen shown below is used to enter the filename in which the data will be stored. If the input is **not** to be saved, only the enter (↵) key should be depressed. Valid EXSYSP filenames are described below in the section on managing user files. If the filename entered currently exists, the user is given the option to specify that it be overwritten with the current data or to rename the file into which the current data are to be saved.



After entering a filename (or choosing not to), the user can specify whether to exit the program at this time or to return to the point where the Quit or <Q> command was entered.

```
Exit Program Y/N (Default=Y):
```

Other OWLKNEST Operating Features

The description of how to operate OWLKNEST concludes with three sections that describe additional user activities that are possible when using OWLKNEST. These activities are:

- Using Technical Information Sheets,
- Managing OWLKNEST user files, and
- Handling system errors

Specific examples of how OWLKNEST can be used are given in the next part of the handbook.

Using Technical Information Sheets

After a technique has been selected, the user may need additional information about the technique in order to make a determination of whether the technique should be further considered for usage. The Technical Information Sheets (TIS) are brief, one page descriptions of the workload techniques included in OWLKNEST and are contained in Appendix A. The TIS are also available on the computer and accessible from the operating system. The TIS are not available during the operation of the OWLKNEST system.

Table 6-3 shows the filenames for each TIS. All these files have an extension of TIS. In order to display a TIS, an OWLKNEST session within EXSYSP must be terminated using either the Quit or <Q> commands as described above. When the DOS prompt (A> or C>) is displayed, enter the following command:

NOTE:	If running OWLKNEST from 360Kb diskettes, Disk 3 contains the OWLKNEST Technical Information Sheets and must be inserted into drive A.

1. Enter **TYPE *FILENAME*.TIS**↵ where

   FILENAME is the name of the desired TIS.

The TIS will be displayed on the screen.

The files containing the TIS may also be sent to the printer or incorporated into word processing documents.

Managing OWLKNEST User Files

At any point during the operation of OWLKNEST, the OWLKNEST user will be able to save the input data. The procedure for saving input data and for subsequently retrieving that data are described above. When saved in accordance with the guidance provided by OWLKNEST, those data are automatically stored as user files either in the OWLKNEST subdirectory (if OWLKNEST is being run from a hard disk) or on Disk 1 in the default drive A (if OWLKNEST is being run from dual floppy diskettes).

EXSYSP filenames

NOTE:	When specifying a filename for storing current inputs, **DO NOT** use **OWLKNEST** as the filename. It will destroy the OWLKNEST knowledge base.

OWLKNEST user files are a convenient mechanism for storing parameters associated with a particular application problem. A file

Table 6-3. OWLKNEST Technical Information Sheet Filenames

Technique	Filename
AHP	AHP.TIS
Bedford	BEDFORD.TIS
Blink Rate	BLINKRAT.TIS
Choice RT Secondary Tasks	CHOICERT.TIS
Closed Questionnaires	CLQUEST.TIS
Comparability Analysis	COMPANLY.TIS
Delphi Interviews	DINTVIEW.TIS
Embedded Secondary Tasks	EMBSEC.TIS
Evoked Potentials	EVOKPOT.TIS
Eye Movement	EYEMOVE.TIS
Heart Rate	HRTRATE.TIS
Heart Rate Variability	HRVAL.TIS
Human Operator Simulator	HOS.TIS
McCracken-Aldrich Task Analysis	MCCALD.TIS
MicroSaint	MICSAINT.TIS
Modified Cooper-Harper	MCH.TIS
Open Ended Questionnaires	OEQUEST.TIS
OW	OW.TIS
Prospective OW	OW.TIS
Prospective SWAT	SWAT.TIS
Prospective TLX	TLX.TIS
Pupil Measures	PUPMEAS.TIS
SIMWAM	SIMWAM.TIS
Sternberg Memory Secondary Tasks	STEINBRG.TIS
Structured Interviews	SINTVIEW.TIS
SWAT	SWAT.TIS
TAWL	TAWL.TIS
Time Estimation Secondary Tasks	TIMEEST.TIS
TLX	TLX.TIS
Tr/Ta Task Analysis	TRTA.TIS
Type 1 Primary Measures	TYPE1.TIS
Type 2 Primary Measures	TYPE2.TIS
Unstructured Interviews	UNITVIEW.TIS
Zaklad/Zachary Task Analysis	ZAKZACH.TIS

containing the current parameters can be created and saved at any point during and at the termination of an OWLKNEST session. This file can then be accessed during a subsequent session with OWLKNEST so that the user can avoid reentering all the parameters associated with that problem.

Valid EXSYSP filenames contain 1 to 8 alphanumeric characters with no special characters. The DOS extension of 3 characters following the period is not permitted. If the default drive is not used, then the filename must be preceded with the drive designator including colon. For example, if the filename, MYFILE, is located on drive b, then the filename in response to EXSYSP prompts would be B:MYFILE.

Viewing OWLKNEST user files

The user may wish to view the data files saved from previous sessions before beginning a new session. If the user wishes to retrieve a previously saved file, the exact name of that file will have to be entered in response to the prompt given with the recover data screen. If the user wishes to save the data entered during any subsequent session with OWLKNEST, the name assigned to the file containing that new data must be different from any used for previously saved data or the new data will overwrite the old data.

To view the OWLKNEST user files, it is necessary to use the "DIR" command while in DOS. If OWLKNEST is being run from a hard disk, the user will enter DIR when the OWLKNEST subdirectory prompt is displayed. The subdirectory list will show 50 files which have a 3-character file extension as part of their respective file names. If OWLKNEST is being run from dual floppy diskettes, the user will enter DIR when the default drive A prompt is displayed. The drive A directory will list 12 files on Disk 1 that have a 3-character file extension as part of their respective file names. In either case, the file names with extensions are used in executing OWLKNEST. All user created files of input data will *not* have a file extension.

Deleting OWLKNEST user files

Once a user file is no longer needed, it may be deleted from the OWLKNEST subdirectory of the hard disk or from Disk 1 in the default drive A. First, the user should display the appropriate directory. The procedure is to type the DOS command ERASE or DELETE, followed by a space, and the name of the user filename that is to be deleted. Then, pressing the enter key causes the named user file to be deleted from the directory.

NOTE:	The user must be sure not to delete any files that have a 3-character name extension. Deleting those types of files will interfere with future attempts to operate OWLKNEST.

Handling System Errors

This section describes system errors that may be encountered during the OWLKNEST session. System errors are those that are related to the overall use of OWLKNEST and the EXSYSP expert system shell. Data entry and usage errors are described in previous sections.

Out of disk space

NOTE:	All diskettes must have been previously formatted in order to be accessed by OWLKNEST.

The OUT OF DISK SPACE error occurs when the user indicates that information is to be stored in a file — e.g., the user inputs are to be saved as a file for subsequent use — and insufficient space is available on the indicated disk. This is most likely to occur when running OWLKNEST from floppy diskette(s). EXSYSP will allow the user to change to a new diskette and then continue with the file saving process.

Out of memory

The PROGRAM TOO BIG TO FIT IN MEMORY and OUT OF MEMORY error messages are generated when the user attempts to rerun OWLKNEST in a single session and insufficient memory is available. The easiest resolution of this problem is to reboot the computer and then rerun OWLKNEST.

PART SEVEN: Sample Problems

To illustrate the use of OWLKNEST, three representative applications are described below:

- Case 1: Early System Design,
- Case 2: Test and Evaluation, and
- Case 3: Preplanned Product Improvement (P^3I).

In addition, this part of HOOT also discusses and illustrates the use of OWLKNEST to perform excursions or what-if types of analyses on previously conducted applications of the software.

The circumstances of each sample case and the what-if example are presented along with user inputs and OWLKNEST recommendations. The numeric values shown for each technique have the following interpretations:

7-8 High recommendation

4-6 Average recommendation

2-3 Low recommendation

The higher the rating values, the more appropriate the technique is to the particular situation. Interpretations of the OWLKNEST ratings and the procedures for obtaining Technical Information Sheets are more fully described in Part Six.

Case 1: Early System Design

The first case illustrates an early system design study with the following set of conditions:

- Only paper specifications exists and no mockup or prototype is available,
- A general idea of how the tasks should be accomplished has been determined,

- Subject matter experts are available with experience on a similar system but access to the comparable system hardware is not possible,
- No more than a week is available for the workload study including preparation and analysis,
- The primary objective is to obtain global workload measures, and
- An easy-to-use technique is preferred, particularly in the areas of preparation and analysis.

Before starting a dialogue with OWLKNEST, the user should organize and summarize what is known about the particular workload problem. The OWLKNEST checklist can assist the user in this regard; it permits the user to predetermine answers to questions that may be posed by OWLKNEST.

Table 7-1 illustrates how a hypothetical user might use a copy of the checklist. The user's knowledge about the problem are shown as check marks (☑). They indicate the following types of input data given as conditions in the description of the problem:

- The availability of operators, hardware systems, subject matter experts, comparable systems, and task data;
- The existence of time constraints; and
- The requirement for ease of use, diagnosticity, and sensitivity.

Furthermore, we might assume that the hypothetical user anticipates and selects the most inclusive response alternatives for questions addressing techniques, performance measures, workload dimensions, and desired output. Finally, let us assume the user has ready access to three types of special equipment for assessing workload.

Note that when using a copy of the OWLKNEST checklist, the user need not predetermine an answer to all 27 of the questions OWLKNEST could pose. Some potential questions address issues that are not anticipated to be relevant to the problem or are simply not of interest to the user.

Table 7-1. OWLKNEST Input for Case 1 — Early System Design

Techniques:	☐ Analytical	☐ Empirical	☑ Both	
Operator:	☐ Available	☑ Not Available		
System/Hardware:	☐ Available	☑ Not Available		
Subject Matter Experts:	☑ Available	☐ Not Available		
Comparable System:	☐ Available	☑ Not Available		
Time Constraints:	☐ None (or one of the following:)			
Preparation:	☑ < 1 day	☐ < 1 week	☐ < 1 month	☐ > 1 month
Data collection:	☐ < 1/2 hour	☐ < 1 hour	☐ < 4 hours	☐ > 4 hours
Operator:	☐ < 1 minute	☐ < 5 minutes	☐ < 15 min	☐ > 15 minutes
Scoring:	☐ < 1 minute	☐ < 5 minutes	☐ < 15 min	☐ > 15 minutes
Data Analysis:	☑ < 1 day	☐ < 1 week	☐ < 1 month	☐ > 1 month
Ease of Use	☑ Preparation ☐ Not an issue	☐ Collection	☐ Scoring	☑ Analysis
Performance:	☐ System	☐ Operator	☑ Both	
Spare Capacity Analysis:	☐ Yes	☐ No		
Available Task Data:	☑ Descriptions ☐ Estimated time of detailed tasks	☑ General task times		
Available Equipment	☑ Audio Tape ☐ EEG	☑ Video Tape ☐ Oculometer	☐ Pupil Diameter ☑ IBM PC	☐ EKG ☐ None
Workload Dimensions:	☐ Auditory ☐ Stress	☐ Cognitive ☐ Time	☐ Motor ☐ Visual	☐ Physical ☑ None
Operator Contact:	☐ Face-to-face	☐ Remote	☐ None	
Real-Time Application:	☐ Yes	☐ No		
Environment:	☐ Operational	☐ Laboratory		☐ Both
Training Time:	☐ < 15 mins	☑ < 1 hour	☐ < 4 hours	☐ > 4 hours
Operator Interference:	☐ None	☐ Minimal	☐ Not a concern	
Operator Intrusiveness:	☐ None permitted ☐ Limited Eye	☐ Limited head ☐ Not a concern	☐ Minimal	
Desired Outputs:	☐ Quantitative	☐ Qualitative		☑ Either
Diagnosticity:	☑ Global	☐ Detailed	☐ Either	
Result anonymity:	☐ Yes	☐ Not a concern		
Sensitivity:	☑ Large	☐ Both subtle and large		

After OWLKNEST is started and the user-OWLKNEST dialogue has run its course, some anticipated questions may not have been asked and some non-anticipated questions may have been asked. In this example, OWLKNEST will *not* query the user about the desired measure of performance or the available equipment. OWLKNEST will, however, ask the user to specify the type of contact possible between the data collector and the subject matter expert, as well as the need to keep the source of data confidential, both queries that our hypothetical user did not anticipate. (In the results given below it is presumed that the user answered these two unanticipated questions with the 'face-to-face' and 'not a concern' options, respectively.)

The point to be stressed is that OWLKNEST will only ask for the input data it needs to evaluate the rules which are activated. Depending upon the user's answers to successive questions, not all possible rules will be activated and those activated may not be activated in the same order.

Case 1 results

Based upon the responses selected for this situation, OWLKNEST ruled out empirical techniques since a representative man-in-the-loop system would not be available for the workload study. The following recommendations would be made for analytical techniques:

Choice		Value
1.	Analytical	8
2.	Expert Opinions	8
3.	Questionnaires	8
4.	*Closed Questionnaires	8
5.	Interviews	6
6.	*Open-ended Questionnaires	6
7.	Prospective Ratings	6
8.	Pro-Global	5
9.	*Pro-OW	5
10.	Pro-Multidimensional	5
11.	*Pro-TLX	5
12.	*Pro-SWAT	5

Note that the choices or techniques recommended represent all the nodes in the technique hierarchy that are assigned non-zero values in the "Then-

conclusions" of activated rules. Actual techniques, represented as the lowest level nodes in the hierarchy, are shown with asterisks.

Interpreting the Case 1 results

While examining these results, the user may wish to know why these techniques have been recommended for assessing workload in this particular early system design study. The reasoning process used in OWLKNEST may, among other things, help the user to better understand the relevant issues involved in this particular study and to better interpret these results.

For example, the user may wish to know why the technique of Closed Questionnaires is rated higher than the technique of Open-ended Questionnaires. If, while viewing the Results screen, the user types in the line number for Closed Questionnaires (i.e., **4**) and presses **ENTER** (↵), OWLKNEST will show that for this study the Closed Questionnaire technique was addressed only in the Then-conclusion of Rule 11. The explanation screen shows that this rule assigns the value of 8 to both types of questionnaire techniques because the relevant higher nodes in the technique hierarchy (analytical, expert opinion and questionnaire techniques, at Levels 1, 2, and 3, respectively) are applicable to the situation under consideration.

For this same results screen, if the user enters the line number for Open-ended Questionnaires (i.e., **6**), OWLKNEST will show that in this study Rules 11 and 136 were both applied and that they assigned values of 8 and 5, respectively, to the technique. Rule 136 assigns a lower value to the technique because the user has indicated that the time to analyze the results obtained in the study is limited to less than one day. OWLKNEST assigns a value of 5 since the analysis of results obtained with Open-ended Questionnaires generally require more time than that.

When more than one rule assigns a value to a particular technique, OWLKNEST computes the average of all non-zero values assigned by the rules. In this case, OWLKNEST adds the values 8 and 5, and divides the

sum of 13 by 2. The resulting truncated value of 6 is assigned to the Open-ended Questionnaire technique for this particular workload study.

To continue this line of inquiry, the user may wish to know the reasoning process which was used by OWLKNEST to eliminate certain techniques from the list of recommendations. First, the user must enter **A** while viewing the Results screen to display **all** the techniques that were activated as a result of the user-OWLKNEST dialogue. If the user does so for this example, OWLKNEST would show that in addition to the 12 techniques that were recommended, 16 other nodes in the technique hierarchy had been activated but subsequently assigned a value of zero and eliminated from further consideration.

One of the eliminated nodes represented **all** empirical techniques. The entire category of empirical techniques was eliminated from further consideration since the user had indicated that neither system operators nor the system of interest were available for the study. The other 15 techniques, all analytical techniques, were eliminated from further consideration because of specified characteristics and requirements of this early design study.

For example, the user could note that the technique of Comparability Analysis was at one point considered a viable technique (because it is an analytical technique) but that it subsequently was eliminated from further consideration. If the user enters the line number of this technique as it is shown on the Results screen that lists all techniques considered, OWLKNEST will display a series of five rules (8, 52, 59, 113, and 118), each of which assigns a value to Comparability Analysis in its conclusions. The technique was eliminated from further consideration in this study by both Rules 52 and 59. Rule 52 assigned a value of zero to Comparability Analysis because the user had indicated that no comparable system was available. Rule 59 also assigned a value of zero because, even if a comparable system were available, the time available to analyze the data was not sufficient.

The examples given here for obtaining explanations for the recommendations provided by OWLKNEST also emphasize a very

important point that the user must remember. Namely, that to adequately interpret the recommendations, the user of OWLKNEST must ultimately rely on his or her own understanding of the unique characteristics and requirements of the workload assessment environment and of the techniques included in this workload assessment data base. OWLKNEST is an aid to decision making, it is not the decision maker.

Case 2: Test and Evaluation

The second case represents the type of situation that could occur when the system has reached a more mature stage in its development. Consider a workload study which has the following characteristics:

- A system prototype and representative operators will be available to the workload analyst for one-half day prior to, during, and for a few minutes following some form of operational field test,
- Detailed descriptions and estimated timing of tasks are available for all operator tasks,
- One month is available for the workload study with about a week each for preparation and analysis, and
- The primary goal of the workload study is to diagnose the sources of overall workload.

Table 7-2 illustrates the input data selected for this problem with the check marks (✔) indicating the selected options. In this case, the user felt that empirical techniques would be best suited and therefore indicated that only empirical techniques should be considered. Eliminating analytical techniques from consideration reduces the number of questions that are posed to the user and speeds up the process.

Case 2 results

For this case, OWLKNEST recommendations would include the following:

Choice	Value
Empirical	8
Subjective	8
Ratings	8
Other Subjective Ratings	8
Questionnaires	8
*Open-ended Questionnaires	8
*Closed Questionnaires	8
Primary	8
Multi-Dimensional Ratings	8
*TLX	8
*SWAT	8
Interviews	6
*Structured Interviews	6
Secondary	6
*Embedded	6
*Unstructured Interviews	5

The numerical value assigned to the interview techniques have been lowered because the time required to analyze the data they produce is generally longer than the time available; unstructured interviews have the lowest value because they are less well defined and harder to conduct successfully than structured interviews. The rating values of the secondary task techniques have been lowered because they also typically require more time than available for analysis. In addition, it generally takes more time than is available to set up and prepare secondary task techniques in an operational setting.

Case 3: Preplanned Product Improvement (P^3I)

The third case represents a study of an existing system for which there is to be a Preplanned Product Improvement (P^3I). The study is to evaluate the P^3I and provide information to the system developers about the advantages and disadvantages of the alternative design from a workload perspective. The workload study has the following set of conditions:

- Operators of the existing system, who also are knowledgeable about characteristics of the improved product, are available,
- No system providing capabilities equivalent to those proposed for P^3I is available,

Table 7- 2. OWLKNEST Input for Case 2 — Test and Evaluation

Techniques:	❑ Analytical ☑ Empirical ❑ Both
Operator:	☑ Available ❑ Not Available
System/Hardware:	☑ Available ❑ Not Available
Subject Matter Experts:	❑ Available ❑ Not Available
Comparable System:	❑ Available ❑ Not Available
Time Constraints:	❑ None (or one of the following:)
Preparation:	❑ < 1 day ☑ < 1 week ❑ < 1 month ❑ > 1 month
Data collection:	❑ < 1/2 hour ❑ < 1 hour ❑ < 4 hours ❑ > 4 hours
Operator:	❑ < 1 minute ❑ < 5 minutes ❑ < 15 min ❑ > 15 minutes
Scoring:	❑ < 1 minute ❑ < 5 minutes ❑ < 15 min ❑ > 15 minutes
Data Analysis:	❑ < 1 day ☑ < 1 week ❑ < 1 month ❑ > 1 month
Ease of Use	❑ Preparation ❑ Collection ❑ Scoring ❑ Analysis ☑ Not an issue
Performance:	❑ System ❑ Operator ☑ Both
Spare Capacity Analysis:	❑ Yes ☑ No
Available Task Data:	❑ Descriptions ❑ General task times ❑ Estimated time of detailed tasks
Available Equipment	☑ Audio Tape ☑ Video Tape ❑ Pupil Diameter ❑ EKG ❑ EEG ❑ Oculometer ☑ IBM PC ❑ None
Workload Dimensions:	❑ Auditory ❑ Cognitive ❑ Motor ❑ Physical ❑ Stress ❑ Time ❑ Visual ☑ None
Operator Contact:	☑ Face-to-face ❑ Remote ❑ None
Real-Time Application:	☑ Yes ❑ No
Environment:	☑ Operational ❑ Laboratory ❑ Both
Training Time:	❑ < 15 mins ❑ < 1 hour ☑ < 4 hours ❑ > 4 hours
Operator Interference:	❑ None ❑ Minimal ☑ Not a concern
Operator Intrusiveness:	❑ None permitted ❑ Limited head ❑ Minimal ❑ Limited Eye ☑ Not a concern
Desired Outputs:	❑ Quantitative ❑ Qualitative ☑ Either
Diagnosticity:	❑ Global ☑ Detailed ❑ Either
Result anonymity:	❑ Yes ☑ Not a concern
Sensitivity:	❑ Large ☑ Both subtle and large

- Detailed descriptions and estimated timing of the operator tasks are available,
- Both time and visual workload dimensions are to be analyzed, and
- Detailed diagnostic information about the source of workload is to obtained.

Table 7-3 illustrates the input data selected for this problem with the check marks (✔) indicating the selected options. For the initial run, no time constraints or ease of use considerations will be indicated.

Case 3 results

For Case 3, OWLKNEST recommendations would include the following:

Choice		Value
1.	Analytical	8
2.	Expert Opinions	8
3.	Prospective Ratings	8
4.	Task Analysis	8
5.	Simulations	8
6.	*Human Operator Simulator	6
7.	Interviews	6
8.	*Structured Interviews	6
9.	*Unstructured Interviews	6
10.	*Delphi Interviews	6
11.	Questionnaires	6
12.	Open-ended Questionnaires	6
13.	*Closed Questionnaires	6
14.	Pro-Multidimensional	6
15.	Task-Based Task Analysis	6
16.	*MicroSaint	6
17.	*TAWL	6
18.	*SIMWAM	6
19.	*Pro-TLX	5
20.	*Pro-SWAT	5
21.	*McCracken-Aldrich	5
22.	*Zaklad-Zachary	5

Table 7-3. OWLKNEST Input for Case 3 — PIP

Techniques:	☐ Analytical ☐ Empirical ☑ Both
Operator:	☑ Available ☐ Not Available
System/Hardware:	☐ Available ☑ Not Available
Subject Matter Experts:	☑ Available ☐ Not Available
Comparable System:	☐ Available ☑ Not Available
Time Constraints:	☑ None (or one of the following:)
Preparation:	☐ < 1 day ☐ < 1 week ☐ < 1 month ☐ > 1 month
Data collection:	☐ < 1/2 hour ☐ < 1 hour ☐ < 4 hours ☐ > 4 hours
Operator:	☐ < 1 minute ☐ < 5 minutes ☐ < 15 min ☐ > 15 minutes
Scoring:	☐ < 1 minute ☐ < 5 minutes ☐ < 15 min ☐ > 15 minutes
Data Analysis:	☐ < 1 day ☐ < 1 week ☐ < 1 month ☐ > 1 month
Ease of Use	☐ Preparation ☐ Collection ☐ Scoring ☐ Analysis ☑ Not an issue
Performance:	☐ System ☐ Operator ☐ Both
Spare Capacity Analysis:	☐ Yes ☐ No
Available Task Data:	☑ Descriptions ☐ General task times ☑ Estimated time of detailed tasks
Available Equipment	☑ Audio Tape ☑ Video Tape ☐ Pupil Diameter ☐ EKG ☐ EEG ☐ Oculometer ☑ IBM PC ☐ None
Workload Dimensions:	☐ Auditory ☐ Cognitive ☐ Motor ☐ Physical ☐ Stress ☑ Time ☑ Visual ☐ None
Operator Contact:	☑ Face-to-face ☐ Remote ☐ None
Real-Time Application:	☐ Yes ☐ No
Environment:	☐ Operational ☐ Laboratory ☐ Both
Training Time:	☐ < 15 mins ☐ < 1 hour ☐ < 4 hours ☑ > 4 hours
Operator Interference:	☐ None ☐ Minimal ☐ Not a concern
Operator Intrusiveness:	☐ None permitted ☐ Limited head ☐ Minimal ☐ Limited Eye ☐ Not a concern
Desired Outputs:	☐ Quantitative ☐ Qualitative ☑ Either
Diagnosticity:	☐ Global ☑ Detailed ☐ Either
Result anonymity:	☐ Yes ☑ Not a concern
Sensitivity:	☐ Large ☑ Both subtle and large

Empirical techniques were eliminated since a system is not available for the operators to utilize during the workload study. Since time constraints were not specified in this run,all task analysis and simulations workload techniques are included in the recommendations. However, interviews and questionnaires techniques as well as the task analysis and simulation techniques (except HOS) have had their values lowered because they do not directly provide information about the desired time and visual workload dimensions. The specified prospective rating and task-based task analysis techniques have lowered values because they are less well defined that the other recommended techniques.

What-If Analysis

OWLKNEST provides the capability to do what-if type analyses. The user can change one or more inputs and determine what the effect would be on the recommended list of techniques. This process is described in more detail in 'Examining the Impact of Alternative Data' section of Part Six. OWLKNEST can be used several times for the same application by changing one or more of the responses given. For example, in the first run, the analyst may choose to respond that no special equipment is available and obtain results based on that answer as one of the inputs. In the next run, however, the analyst may want to see what other techniques would be appropriate if audio and video recording equipment were available. In this case, the suggested list might include different techniques. In this way, the analyst will be provided with some information on which to base decisions as to whether additional resources should be allocated to the workload assessment effort.

The system also is intended to be used iteratively across time to address specific circumstances facing the user at different points in system development. For example, a workload analysis mav he desired early in system design. At this early point, no detailed information will be available and gross predictions of workload will be sought using analytical techniques. Later, after the initial development is complete and a prototype

is available, OWLKNEST might be used again to suggest workload techniques based on the currently available information and resources. At this point, empirical techniques will be feasible due to the prototype and the larger body of system information. Hence, OWLKNEST can be fruitfully used throughout the development cycle of the system.

What-If Analysis Example

Assume that after running OWLKNEST for the P^3I case above, new conditions and requirements were imposed upon the workload analyst. First, it was determined that the current system would be modified and made available for an operational test that would include a workload analysis. Additionally, the analyst was informed that only about a month could be allocated for the study. Therefore, the analyst decided to allocate a week each to preparation for the study and analysis of the data. Also, in part because the abbreviated time available for the study was insufficient for applying task analysis and simulation techniques, the analyst decided that only empirical techniques were to be considered.

Since the initial P^3I input parameters had been saved in a file, the analyst began this new run by indicating that those values were to be recovered. Then, the analyst selected the **C** - Change option of the old Results screen, and the input parameters were modified in accordance with the new information. Table 7-4 shows the input data for this example. As illustrated by the "change" notations given in the left margin of the table, the input was changed for five of the previously asked questions. Then, when the **R** - Rerun option of the Results screen was selected, OWLKNEST proceeded to display eight additional questions that were not previously asked (shown by the "new" notations in the left margin of Table 7-4).

Results of the what-if analysis

As a result of the input data, OWLKNEST produced the following set of recommended workload assessment techniques:

Table 7-4. OWLKNEST Input for the What-if Analysis Example

Change	Techniques:	☐ Analytical ☑ Empirical ☐ Both
	Operator:	☑ Available ☐ Not Available
Change	System/Hardware:	☑ Available ☐ Not Available
	Subject Matter Experts:	☑ Available ☐ Not Available
	Comparable System:	☑ Available ☐ Not Available
Change	Time Constraints:	☐ None (or one of the following:)
- New	Preparation:	☐ < 1 day ☑ < 1 week ☐ < 1 month ☐ > 1 month
	Data collection:	☐ < 1/2 hour ☐ < 1 hour ☐ < 4 hours ☐ > 4 hours
	Operator:	☐ < 1 minute ☐ < 5 minutes ☐ < 15 min ☐ > 15 minutes
	Scoring:	☐ < 1 minute ☐ < 5 minutes ☐ < 15 min ☐ > 15 minutes
- New	Data Analysis:	☐ < 1 day ☑ < 1 week ☐ < 1 month ☐ > 1 month
	Ease of Use	☐ Preparation ☐ Collection ☐ Scoring ☐ Analysis ☑ Not an issue
- New	Performance:	☐ System ☐ Operator ☑ Both
- New	Spare Capacity Analysis:	☐ Yes ☑ No
	Available Task Data:	☑ Descriptions ☐ General task times ☑ Estimated time of detailed tasks
	Available Equipment	☑ Audio Tape ☑ Video Tape ☐ Pupil Diameter ☐ EKG ☐ EEG ☐ Oculometer ☑ IBM PC ☐ None
	Workload Dimensions:	☐ Auditory ☐ Cognitive ☐ Motor ☐ Physical ☐ Stress ☑ Time ☑ Visual ☐ None
	Operator Contact:	☑ Face-to-face ☐ Remote ☐ None
- New	Real-Time Application:	☑ Yes ☐ No
- New	Environment:	☐ Operational ☐ Laboratory ☐ Both
Change	Training Time:	☐ < 15 mins ☐ < 1 hour ☑ < 4 hours ☐ > 4 hours
- New	Operator Interference:	☐ None ☐ Minimal ☑ Not a concern
- New	Operator Intrusiveness:	☐ None permitted ☐ Limited head ☐ Minimal ☐ Limited Eye ☑ Not a concern
	Desired Outputs:	☐ Quantitative ☐ Qualitative ☑ Either
	Diagnosticity:	☐ Global ☑ Detailed ☐ Either
	Result anonymity:	☐ Yes ☑ Not a concern
	Sensitivity:	☐ Large ☑ Both subtle and large

Choice		Value	
		Current	Previous
1.	Empirical	8	0
2.	Subjective	8	None
3.	Ratings	8	None
4.	Interviews	6	6
5.	*Structured Interviews	6	6
6.	Questionnaires	6	6
7.	*Open-ended Questionnaires	6	6
8.	*Closed Questionnaires	6	6
9.	Primary	6	None
10.	Multi-dimensional Ratings	6	None
11.	*TLX	6	None
12.	*SWAT	6	None
13.	Other Subjective Ratings	6	None
14.	*Unstructured Interviews	5	None
15.	*Type 2	5	None
16.	Secondary	5	None
17.	*Embedded	5	None
18.	Analytical	0	8

Note that two set of rating values are displayed for each technique listed, one determined by the current "what-if" input data and the other determined by inputs previously entered for the original P^3I problem.

The most obvious difference between the previous and current sets of recommendations is that the former were all analytical and the current are all empirical. Since interview techniques and questionnaires can be used in either case, these techniques were equally appropriate in both. It should be noted that while the general categories of empirical, subjective, and rating techniques were given the highest possible rating values for the current case, no actual workload assessment technique was assigned such a high value. Values were reduced because of the time constraints imposed upon the study and because most of the techniques do not directly provide information about specific workload dimensions.

The results obtained from this what-if example could also be saved and utilized in subsequent sessions with OWLKNEST. By appropriate modification of the input parameters, a more narrowly focused set of recommendations may be obtained. Typically, the number of details

supplied by the user is inversely related to the quantity of recommendations obtained. That is, the more specific the inputs, the smaller the output set.

As has been emphasized elsewhere in this handbook, the user must make the final determination of which techniques should be further pursued and ultimately used in a workload study. The information provided in the TIS for each of the techniques should provide valuable assistance in this determination.

PART EIGHT: Conclusions

The Army's increasing reliance on complex high technology systems has changed the role of the human who operates these systems. In general, new and emerging military systems impose requirements on the operator for more mental or cognitive capabilities than was true for systems in the past.

Army regulations require that the design, development, and evaluation of new systems must proceed in a manner which will ensure that the systems do not cause the demand for mental skills to exceed the capabilities of targeted system operators. If this principal is violated, and if the system demands on an operator are greater than the operator's capability to respond, the system may not perform to standards. Therefore, it follows that an effective operator workload (OWL) assessment program must be a component of the materiel acquisition process.

In addition, an effective OWL assessment program is also critical for optimizing the outcome of the following activities:

- o Allocate workload-imposing functions and tasks among operator personnel, hardware, and software components of systems,
- o Design organizational units and develop operator, crew, and small unit tactics, techniques, and procedures in a manner that will address workload issues and concerns, and
- o Establish procedures for the selection, classification, and training of operators and small units to effectively manage workload in operational situations.

Fortunately, a variety of OWL assessment techniques are available and many are well documented in published papers and reports. Unfortunately, it is difficult for most workload analysts to readily determine which techniques are most appropriate for their particular workload study. It was in this light that OWLKNEST was created to aid the workload analyst involved in the development of Army systems, organizations, doctrine, and training programs.

Those who have a need to predict or evaluate operator workload will find expert advice at their fingertips when they use OWLKNEST. It will recommend and provide information on appropriate OWL assessment techniques based on the user's specification of the characteristics of the particular workload assessment problem and of the available resources. OWLKNEST is applicable across all phases of the materiel acquisition process and all functions and components of the force integration model. It is a comprehensive, easy-to-use tool which emphasizes techniques suitable for operational and field application as well as for more traditional analytical environments.

Although originally designed for application to military problems, there is a clear opportunity to transfer the OWLKNEST technology to more general industrial and commercial applications. Although the terms may be different, concern with an operator's ability to control systems that use complex technology is the same. Similarly, the workload techniques are not application specific, but can be used in a variety of domains. OWLKNEST can provide the information and guidance necessary to select appropriate workload techniques for a broad range of applications.

As well as providing recommendations for workload assessment techniques, OWLKNEST also is envisioned to serve as a clearinghouse of knowledge for workload concepts, assessment methodologies, and technique applications. OWLKNEST serves as an aid for organizing information relevant to the entire domain of workload issues and concerns. This capability of the tool is the basis for maintaining the currency and relevance of the OWLKNEST knowledge base and for ultimately increasing its applicability to areas other than those that are strictly military and operator centered.

Refinements and enhancements of the OWLKNEST knowledge base can and should continue as more information and experience with the current set of techniques is obtained, and as other techniques are identified for inclusion. The guidance provided by OWLKNEST can and should be further validated in its future applications.

OWLKNEST Survey

The most important information source to steer the OWLKNEST refinement process is feedback from its users. For this purpose, the OWLKNEST User Survey has been developed and is contained in Appendix B. OWLKNEST users are encouraged to complete copies of this survey as they engage in successive applications of the tool. The completed survey forms should be returned as indicated. If resources necessary to maintain a current knowledge base and to otherwise enhance OWLKNEST are made available, the users' input will be incorporated into future versions of this expert system. Likewise, user feedback will aid in the development of other expert system tools that might be developed to provide information and guidance for the selection of other types of measurement techniques.

REFERENCES

Boehm-Davis, D. A. (Ed.). (1988). Expert systems [Special issue]. *Human Factors*, 30(4), 377-443.

Byers, J. C., Hill, S. G., Zaklad, A. L., & Christ, R. E. (1989). *Aquila system report* (TR 2075-4a). Willow Grove, PA: Analytics, Inc.

Casper, P. A., Shively, R. J., & Hart, S. G. (1986). Workload Consultant: A microprocessor-based system for selecting workload assessment procedures. In *Proceedings of the IEEE Conference on Systems, Man, and Cybernetics* (pp. 1054-1059). Piscataway, NJ: IEEE Service Center.

Casper, P. A., Shively, R. J., & Hart, S. G. (1987). Decision support for workload assessment: Introducing WC FIELDE. In *Proceedings of the Human Factors Society 31st Annual Meeting* (pp 72-76). Santa Monica, CA: Human Factors Society.

Christ, R. E., Bulger, J. P., Hill. S. G., & Zaklad, A. L. (1990). *Incorporating operator workload issues* and concerns into the system acquisition process: A pamphlet for Army managers (Research Product 90-30). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC Assession No. AD A228 489)

Fotta M. E., & Davis J. M. (1988, October). *An introduction to artificial intelligence and expert systems*. Unpublished workshop presented at the 32nd Annual Meeting of the Human Factors Society. Anaheim, CA.

Gopher, D., & Donchin, E. (1986). Workload: An examination of the concept. In K. R. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance: Vol. II. Cognitive processes and performance*. N.Y.: John Wiley & Sons.

Harmon, P., & King, D. (1985). *Expert systems: Artificial intelligence in business*. New York, NY: John Wiley & Sons.

Hill, S. G., Byers, J. C., Christ, R. E., & Zaklad, A. L. (1989). *LOS-F-H system report* (TR 2075-4b). Willow Grove, PA: Analytics, Inc.

Hill, S. G., & Harris, R. M. (1990). OWLKNEST: A knowledge-based expert system for selecting operator workload techniques. In W. Karwowski, A. M. Genaidy, andS. S. Asfour (Eds.), *Computer-aided design: A researcher's guide* (pp. 213-225). London: Taylor and Francis.

Hill, S. G., Iavecchia, H. P., Zaklad, A. L., Christ, R. E., & Sams, M. R. (In process). *Application and validation of workload assessment techniques: Final report* (Technical Report). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Hill, S. G., Lysaght, R. J., Bittner, A. C., Jr., Bulger, J. P., Plamondon, B. O., Linton, P. M., & Dick, A. O. (1987). *Operator workload assessment rrogram for the army: Results from requirements document review and sser interview aanalysis* (DTR 2075-2), Willow Grove, PA: Analytics, Inc.

Iavecchia, H. P., Linton, P. M., Byers, J. C., & Harris, R. M. (1989). *UH-60 system report* (TR 2075-4c). Willow Grove, PA: Analytics, Inc.

Ignizio, P. (1990). *Introduction to expert systems: The development and implementation of rule-based expert systems*. Hightstown, NJ: McGraw-Hill.

John, J. G., Klein, G. A., & Taylor, J. (1986). Comparison-based prediction for front-end analysis. In *Proceedings of the Human Factors Society 30th Annual Meeting* (pp. 149-153). Santa Monica, CA: Human Factors Society.

Lysaght, R. J., Hill, S. G., Dick, A. O., Plamondon, B. D., Linton, P. M., Wierwille, W. W., Zaklad, A. L., Bittner, A. C., Jr., & Wherry, R. J., Jr. (1989). *Operator workload: Comprehensive review and evaluation of operator workload methodologies*, (Technical Report No. 851). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC Assession No. AD A212 879).

Miller, B. H. (1988, September/October). Expert systems: An introduction. *PC AI*, Volume 2, No. 3, pp. 26-29.

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. Boff, L. Kaufman and J. Thomas (Eds.), *Handbook of perception and human performance: Vol II. Cognitive process and performance.* NY: John Wiley & Sons.

Shaffer, M. T., Shafer, J. B., & Kutch, G. B. (1986). Empirical workload and communication: Analysis of scout helicopter exercises. In *Proceedings of the Human Factors Society 30th Annual Meeting* (pp. 628-632). Santa Monica, CA: Human Factors Society.

U.S. Army Soldier Support Center (1986). *Early comparability analysis (ECA) procedural guide.* Washington, DC: Department of the Army.

Waterman, D. A. (1986). *A guide to expert systems.* Reading, MA: Addison-Wesley.

Wierwille, W. W., & Williges, B. C. (1980). *An annotated bibliography on operator mental workload assessment* (SY-27R-80). Patuxent River, MD: Naval Air Test Center.

APPENDIX A

TECHNICAL INFORMATION SHEETS

Analytical Hierarchy Process (AHP)

DESCRIPTION: This is a relative workload assessment technique. All possible pairs of tasks or mission segments are presented to the operator. If one of the pair is judged to have higher workload, the operator is asked to judge by how much (on a scale from 1 to 5, with 1 being equal workload and 5 being extremely high relative workload). This relative workload assessment is most appropriate for post-session ratings.

SENSITIVITY: AHP has been shown to correspond closely with Bedford scale ratings. Was also shown to be reliable and correspond closely with performance. However, insufficient information is currently available to make strong conclusions.

DIAGNOSTICITY: Sufficient information is not available.

INTRUSION: n/a

IMPLEMENTATION REQUIREMENTS:

Data Collection: The ratings can be obtained via paper and pencil, verbally or electronically.

Operator Training: Some practice will be necessary to familiarize the operators with the procedure and the ratings.

OPERATOR ACCEPTANCE: It was successfully used in an application, but sufficient information is not yet available.

SAFETY: n/a

PHYSICAL SPACE REQUIRED: n/a

PORTABILITY: n/a

INTEGRATION INTO SYSTEM: n/a

RESTRICTIONS: n/a

RELATIVE COST OF USE:

Testing time: Depends on the number of paired comparison judgments that need to be made.

Equipment: Whatever is needed for data collection. In addition, computer access is helpful for data reduction and analysis.

Setup and support: Minimal.

Data analysis: Procedures are still being developed. Lidderdale used a "consensus" method for combining subjects' ratings. Vidulich and Tsang used the AHP ratings as data in standard statistical tests. A single data analysis approach has not yet been generally accepted.

COMMENTS:

REFERENCES:

Lidderdale, I.G. (1987). Measurement of aircrew workload during low-level flight . In A. H. Roscoe (Ed.), *The practical assessment of pilot workload, AGARDograph No. 282* (pp. 69-77). Neuilly Sur Seine, France: AGARD.

Lidderdale, I.G., & King, A.H. (1985). *Analysis of subjective ratings using the analytical hierarchy process; A microcomputer program.* High Wycombe, England: OR Branch NFR, HQ STC, RAF

Vidulich, M., & Tsang, P.S. (1987). Absolute magnitude estimation and relative judgement approaches to subjective workload assessment. In *Proceedings of the Human Factors Society 31st Annual Meeting*. Santa Monica, CA: Human Factors Society.

Bedford Scale

DESCRIPTION: Uses a decision tree structure to obtain estimates of workload. It is based on the Cooper-Harper scale and obtains ratings from 1 to 10. It was developed as a means to obtain workload estimates from pilots specifically concerning "spare capacity". It has been used in real-time flight workload estimation. Could potentially be used with visual recreation via video tape.

SENSITIVITY: Has been found to differentiate between flight segments in a military combat aircraft application. Corresponded well to flight segment ratings obtained by the Analytical Hierarchy Process (AHP), although not quite as well to ratings obtained via NASA-TLX and a unidimensional overall workload scale. Insufficient validation research is available to make more conclusive statements.

DIAGNOSTICITY: The rating scale is based on "spare capacity" but does not provide for multidimensional aspects of OWL. It is more of a global scale of workload.

INTRUSION: Little, although it does require a judgment. There was concern (as with most subjective measures) that the judgment might interfere with flight duties, but ratings were able to be obtained real-time.

IMPLEMENTATION REQUIREMENTS:

Data Collection: Some method for collecting the ratings is needed -- either a 10 key pad or communications medium with which the operator can report the rating verbally. A copy of the scale for reference is also useful. For example, Lidderdale (1987) reports using a copy of the scale on the flight suit kneepad.

Operator Training: The operators must be given opportunity to become familiar with the rating scale, therefore some practice is necessary, although the scale is reported to be easy to understand.

OPERATOR ACCEPTANCE: The scale has been reported to be well received by pilots.

SAFETY: Plans must be made as to what to do if the operator is too busy to give a rating. Ratings should be secondary to the primary concern with operational safety (e.g., flying a plane or controlling a land vehicle).

PHYSICAL SPACE REQUIRED: Only that for any data collection device (e.g., a dedicated keypad).

PORTABILITY: n/a

INTEGRATION INTO SYSTEM: n/a

RESTRICTIONS: n/a

RELATIVE COST OF USE:

Testing time: Minimal.

Equipment: Minimal.

Setup and support: Minimal.

Data analysis: Descriptive and inferential statistics can be used. Graphical representations are useful. Caution is advised in assuming an interval scale, therefore non-parametric analysis may be more appropriate.

COMMENTS:

REFERENCES:

Lidderdale, I. G. (1987). Measurement of aircrew workload during low-level flight . In A. H. Roscoe (Ed.), *The practical assessment of pilot workload, AGARDograph No. 282* (pp. 69-77). Neuilly Sur Seine, France: AGARD.

Roscoe, A. H. (1987a). In-flight assessment of workload using pilot ratings and heart rate. In A. H. Roscoe (Ed.), *The practical assessment of pilot workload, AGARDograph No. 282* (pp. 78-82). Neuilly Sur Seine, France: AGARD.

Wainwright, W. (1987). Flight test evaluation of crew workload. In A. H. Roscoe (Ed.), *The practical assessment of pilot workload, AGARDograph No. 282* (pp. 60-68). Neuilly Sur Seine, France: AGARD.

Blink Rate

DESCRIPTION: There are two types of eye blinks — reflex and spontaneous. The spontaneous blink is of interest in the study of workload. It has been demonstrated that blink rate provides a measure directly related to task demands and to task loading (Bauer et al., 1987).

SENSITIVITY: Varies

DIAGNOSTICITY: Varies

INTRUSION: Yes.

IMPLEMENTATION REQUIREMENTS:

Data Collection:

Operator Training:

OPERATOR ACCEPTANCE: Little formal work done on this issue.

SAFETY:

PHYSICAL SPACE REQUIRED: Equipment used to measure eye blinks require some space.

PORTABILITY: The measure is normally used only in a laboratory environment.

INTEGRATION INTO SYSTEM:

RESTRICTIONS: While several techniques exist for measuring eye blinks, most are not easily applicable in an operational environment.

RELATIVE COST OF USE:

Testing time:

Equipment:

Setup and support:

Data analysis:

COMMENTS: Because vision is a major information acquisition sensory system, many investigators have focused efforts on determining how the system functions and acquires information under varying workload conditions. While both eye movements and pupil diameter have received considerable attention in these regards, blink rate has not. There is some indication that measures of eye blinking could be useful, especially in conjunction with other measures of eye behavior. All measures of eye behavior are generally suitable only in a laboratory environment.

REFERENCES:

Bauer, L.O., Goldstein, R., & Stern, J.A. (1987). Effects of information processing demands on physiological response patterns. *Human Factors*, 29, 213-234.

Choice Reaction Time

DESCRIPTION: Choice reaction time has been used as a secondary task to reflect OWL levels on primary tasks. The subject is presented with more than one stimulus and must generate a different response for each one. Visual or auditory stimuli may be employed and the response mode is usually manual. It is theorized that choice reaction time imposes both central-processing and response selection demands.

SENSITIVITY: Has been shown to be sensitive to differences in aircraft task difficulty as defined by mission scenarios which required 21 different flight tasks.

DIAGNOSTICITY: Gives a global measure of workload. It can be examined with respect to specific instances within a scenario to determine the OWL associated with a specific task. However, such a determination requires repeated administration of the Choice RT task in association with the specific task to produce reliable results.

INTRUSION: Little, although there may be situations where Choice RT might interfere with the primary task. Studies have used Choice RT successfully in flight simulators without any interference with flight duties.

IMPLEMENTATION REQUIREMENTS:

Data Collection: Some method is needed to collect the operator's responses and store the data. It is preferable to store the data with some type of time stamp in order to reference the Choice RT responses to specific events.

Operator Training: Operators must be given an opportunity to practice the Choice RT task as well as establish a baseline for their responses to be used for data analysis.

OPERATOR ACCEPTANCE: Pilots in flight simulators have been receptive to performing Choice RT tasks within flight scenarios.

SAFETY: Plans must be made as to what to do if operators are too busy to respond to the Choice RT task. The task should be secondary to the primary concern with operational safety (e.g., flying an operational aircraft or controlling a land vehicle).

PHYSICAL SPACE REQUIRED: n/a

PORTABILITY: Varies as a function of the particular set-up (i.e., hardware and software) for the Choice RT task.

INTEGRATION INTO SYSTEM: Varies as a function of the particular set-up (i.e., hardware and software) for the Choice RT task.

RESTRICTIONS: Varies - see Safety

RELATIVE COST OF USE:

Testing time: A sufficient number of presentations is needed on the Choice RT task to have reliable data for each operator within a test session.

Equipment: Minimal - medium.

Setup and support: Minimal - medium.

Data analysis: Descriptive and inferential statistics can be used on response time scores and error scores. Graphical representations are useful.

REFERENCES:

Bortolussi, M., Hart, S. G., & Shively, R. (1987). Measuring moment-to-moment pilot workload using synchronous presentations of secondary tasks in a motion-based trainer. In *Proceedings of the Fourth Symposium on Aviation Psychology*. Columbus, OH: Ohio State University.

Bortolussi, M., Kantowitz, B. H., & Hart, S. G. (1986). Measuring pilot workload in a motion base trainer. *Applied Ergonomics, 17*, 278-283.

Closed Questionnaires

DESCRIPTION: Questionnaires are forms in which written questions are asked in a fixed order and format and to which respondents write their answers. The questions may be open-ended, allowing respondents to write in their own words and make any answer, or close-ended, where the choice of answers has been previously established, such as multiple choice or true and false. Questionnaires should be used whenever possible to obtain the subtle, detailed information that might not be obtained from rating scales.

SENSITIVITY: Variable.

DIAGNOSTICITY: Detailed information that might not otherwise be obtained can be drawn from interviews.

INTRUSION: Depends on the length of the questionnaire, but for the most part, questionnaires will not be appropriate during real-time operation. Answering questions will require attention and will distract the operator from the primary task.

IMPLEMENTATION REQUIREMENTS:

Data Collection: The most common implementation is via paper and pencil, however, questionnaires can be administered via computer if available.

Operator Training: Minimal.

OPERATOR ACCEPTANCE: In general, questionnaires are well-accepted. However, if questions are not clear, or operators are asked too many questions too often, acceptance may decrease.

SAFETY: Plans must be made as to what to do if the operator is too busy to give a rating. Ratings should be secondary to the primary concern with operational safety (e.g., flying a plane or controlling a land vehicle).

PHYSICAL SPACE REQUIRED: n/a

PORTABILITY: n/a

INTEGRATION INTO SYSTEM: n/a

RESTRICTIONS: n/a

RELATIVE COST OF USE:

Testing time: Can vary in time, depending on how many questions are asked and whether they are open or close-ended.

Equipment: Paper and pencil (or can be computer-based).

Setup and support: Careful development is needed. Little support.

Data analysis: Qualitative and quantitative.

COMMENTS:

REFERENCES:

Dyer, R., Matthews, J., Wright, C., & Yudowitch, K. (1976). *Questionnaire Construction Manual* (TCATA DAHC-19-74-C-0032). Ft. Hood, TX: ARI.

U.S. Army Test and Evaluation Command (1976). *Questionnaire and interview design, Subjective testing techniques* (TECOM Pam 602-1, Vol. I). Aberdeen Proving Ground: USATECOM.

Meister, D. (1985). *Behavioral analysis and measurement methods*. New York: John Wiley and Sons.

Comparability Analysis

DESCRIPTION: Comparability analysis techniques refer to a family of front-end analysis methodologies, each of which involves comparisons between proposed (future) systems and similar (fielded) predecessor systems. These techniques use the physical and functional similarities between the existing and proposed systems to extrapolate data from the fielded system and apply them to the conceptual system. The objective of these techniques is to identify components of the system that are termed "high-drivers" in that they impose inordinate demands on the manpower, personnel, and training (MPT) resources.

SENSITIVITY: The sensitivity of comparability analysis techniques is a function of the reliability and sensitivity of workload data that exists for the predecessor system and the degree of similarity between the predecessor and conceptual system.

DIAGNOSTICITY: The diagnosticity of comparability analysis techniques is a function of the diagnosticity of workload data that exists for the predecessor system and the degree of similarity between the predecessor system and the conceptual system.

INTRUSION: n/a

IMPLEMENTATION REQUIREMENTS:

Data Collection: All comparability analysis techniques require relatively complete operator tasks lists for each system under study. Some also require information which reflects the way the systems are designed, deployed, and operated. Subject matter experts are required to provide system-related information and data and to rate the workload associated with the tasks.

Operator Training: n/a

OPERATOR ACCEPTANCE: n/a

SAFETY: n/a.

PHYSICAL SPACE REQUIRED: n/a

PORTABILITY: n/a

INTEGRATION INTO SYSTEM: n/a

RESTRICTIONS: Currently, comparability analysis is less a well defined technique that it is a generalized procedure. The general procedure offers a fairly straightforward workload prediction technique, but only if workload data are available on a predecessor system. Unfortunately, most operational systems do not have a workload database.

RELATIVE COST OF USE:

Testing time: n/a

Equipment: n/a

Setup and support: The up front cost can be quite high. A thorough task analysis must be performed on both the predecessor and the conceptual systems. Workload data must already exist or be collected for the predecessor system prior to the initiation of the comparability analysis.

Data analysis:

COMMENTS: The more widely recognized comparability techniques include Early Comparability Analysis (ECA), Hardware versus Manpower (HARDMAN) analysis, and the automated version of HARDMAN called Man-Integrated System Technology (MIST) or HARDMAN II.

FUTURE DEVELOPMENTS: A number of efforts are underway to enhance the power of comparability techniques and procedures. These future developments are typically

named in a manner that links them to their predecessors, e.g., Enhanced HARDMAN or HARDMAN III. Some of these future developments contain large databases of task information but none currently in development have a workload database.

REFERENCES:

John, P.G., Klein, G.A., & Taylor, J. (1986). Comparison-based prediction for front-end analysis. In *Proceedings of the Human Factors Society 30th Annual Meeting* (pp. 149-153). Santa Monica, CA: Human Factors Society.

McManus, J.C. (1979). *Equipment comparability techniques used during early system design.* (AFHRL Technical Report 79-24. ADA 071411).

Shaffer, M.T., Shafer, J.B., & Kutch, G.B. (1986). Empirical workload and communication: Analysis of scout helicopter exercises. In *Proceedings of the Human Factors Society 30th Annual Meeting* (pp. 628-632). Santa Monica, CA: Human Factors Society.

U.S. Army Soldier Support Center, National Capital Region (1987). *Early comparability analysis (ECA) procedure guide.* 200 Stovall Street, Alexandria, VA: USASSC-NCR.

U.S. Army TRADOC Analysis Command (1988). *A comparison of three MANPRINT methodologies: Early comparability analysis (ECA), HARDMAN comparability methodology (HCM), and training effectiveness analysis (TEA).* (TRAC-WSMR Pam 602-2). White Sands Missile Range , NM: TRAC-WSMR.

Delphi Interviews

DESCRIPTION: The Delphi method is "... a process whereby subjective judgments or the implicit decision-making processes of experts can be made more objective and explicit:" (Meister, 1985, p. 423). Generally, Delphi is administered to a group of SMEs. The eventual goal is to arrive at a group consensus. The Delphi Technique involves several phases, most of which are iterations or rounds in which the results of previous rounds are summarized and returned with a questionnaire to the group of SMEs. The method is most applicable to situations in which existing referents or comparison systems are not available, or where extrapolation or prediction are required. The validity and reliability of the Delphi method is subject to the same constraints as any other subjective method, but where such methods are required, the more structure Delphi method may strengthen the results.

SENSITIVITY: Variable.

DIAGNOSTICITY: Variable.

INTRUSION:

IMPLEMENTATION REQUIREMENTS:

Data Collection: A multi-phase data collection effort will be required where the problem and ground rules are defined; participating SMEs are identified; initial materials distributed to participants (either by mail or other means such as in a conference); information is obtained, summarized, and put together as a second round of inquiry. The iterative process continues until participants have reached some consensus.

Operator/SME Training: SME training will be accomplished via the instructions written and given to the SMEs.

OPERATOR ACCEPTANCE:

RELATIVE COST OF USE:

Testing time: Variable.

Equipment: Can use paper and pencil or computers to collect data.

Data analysis: Depending on specifics, data can be quantitatively or qualitatively analyzed.

COMMENTS:

REFERENCES:

Meister, D. (1985). *Behavioral analysis and measurement methods.* New York: John Wiley and Sons.

Dalkey, N.C. (1969). *The Delphi method: An experimental study of group opinion..* Rand Corporation, Santa Monica.

Linstone, H.A. and Turoff, M. (Eds.) (1975). *The Delphi method: Techniques and applications.* Reading, MA: Addison-Wesley Publishing Co.

Embedded Secondary Tasks

DESCRIPTION: Embedded secondary task is a technique in which an existing sub-task of a multi-task system is utilized as a secondary task which is fully integrated with existing hardware, software, and the operator's concept of the mission environment. Such embedded secondary tasks are radio communication task, recognition of emergency conditions and navigational problems during simulated flights.

SENSITIVITY: Has been shown to be sensitive to workload in simulated aircraft environments by varying the task difficulty of the embedded secondary task to determine breakdowns in human performance.

DIAGNOSTICITY: Varies as a function of the sub-task which is used as a secondary task.

INTRUSION: Essentially none

IMPLEMENTATION REQUIREMENTS:

Data Collection: Some method is needed to collect the operator's responses and store the data with a time stamp. It is anticipated that existing simulators will be designed for this kind of data collection without any need for further development.

Operator Training: Operators must be given an opportunity to practice on the system such that a baseline can be established for stable performance prior to any task loadings with the embedded secondary task.

OPERATOR ACCEPTANCE: High

SAFETY: The task loadings on the embedded secondary task should be determined so as not to endanger the safety of operators (operational aircraft or controlling a land vehicle).

PHYSICAL SPACE REQUIRED: n/a

PORTABILITY: n/a

INTEGRATION INTO SYSTEM: n/a

RESTRICTIONS: Variable - see Safety

RELATIVE COST OF USE:

Testing time: A sufficient number of presentations is needed on the embedded secondary task to have reliable data for each operator within a test session.

Equipment: existing equipment

Setup and support: existing support

Data analysis: Varies as a function of the embedded secondary task selected, therefore descriptive and inferential statistics may be appropriate. Graphical representations are useful.

COMMENTS:

REFERENCES:

Shingledecker, C. A. (1987). In-flight workload assessment using embedded secondary radio communications tasks. In A. H. Roscoe (Ed.), *The practical assessment of pilot workload, AGARDograph No. 282* (pp. 11-14). Neuilly sur Seine, France: AGARD.

Wierwille, W. W., Casali, J. G., Connor, S. A., & Rahimi, M. (1985). Evaluation of the sensitivity and intrusion of mental workload estimation techniques. In W. Roner (Ed.), *Advances in Man-Machine Systems Research, Volume 2*. (pp. 51-127). Greenwich, CT: J.A.I. Press.

Evoked Potentials

DESCRIPTION: The portion of the brain wave activity that is a response associated with an external stimulus. By performing ensemble averaging across time interval following multiple presentations of the stimulus, the ECP associated with such external stimuli will be enhanced. With respect to operator workload, the positive component of the ECP that occurs approximately 300 msec after stimulus presentation (P300) has been demonstrated to reflect workload levels (e.g., P300 amplitude). The stimulus is usually related in some manner with the task to be performed.

SENSITIVITY: Has been shown to be sensitive to workload levels in controlled laboratory situations. The tasks used in the studies can be characterized as tracking type tasks.

DIAGNOSTICITY: Gives a global measure of workload. However, in controlled laboratory environments there have been demonstrations that the P300 latency measure can be used to isolate the locus of human performance changes, for example, cognitive processing vs. response selection.

INTRUSION: For system evaluation and testing environments, ECP recordings could significantly intrude on operator performance.

IMPLEMENTATION REQUIREMENTS:

Data Collection: Sophisticated hardware and software is needed as well as highly specialized trained personnel.

Operator Training: No special training however operators would need prior exposure to such recording apparatus (e.g., electrodes) before an actual test session.

OPERATOR ACCEPTANCE: In most system evaluation and testing environments, operators would probably feel constrained by such recording techniques.

SAFETY: There are no safety hazards with such recording techniques.

PHYSICAL SPACE REQUIRED: Recording device needed; can be large.

PORTABILITY: Very limited in most situations.

INTEGRATION INTO SYSTEM: Highly unlikely unless in controlled laboratory environments.

RESTRICTIONS: n/a

RELATIVE COST OF USE:

Testing time: A sufficient number of stimulus presentations is needed on the task to have reliable data (ECP) for each operator within a test session.

Equipment: Expensive

Setup and support: Expensive -- highly trained staff is needed.

Data analysis: Descriptive and inferential statistics can be used on P300 amplitude and latency scores. Graphical representations are useful.

COMMENTS:

REFERENCES:

Chapman, R. M., McCrary, J. W., & Chapman, J. A. (1978). Short-term memory: The "storage" component of human brain responses predicts recall. *Science, 202,* 1211-1214.

Donchin, E., Kramer, A. F., & Wickens, C. D. (1986). Applications of brain event-related potentials to problems in engineering psychology. In M. G. H. Coles, E. Donchin, & S. Porges (Eds.), *Psychophysiology: Systems, processes, and applications* (pp. 702-718). New York: Guilford Press.

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition, 11,* 539-550.

Eye Movement

DESCRIPTION: A number of procedures exist for measuring eye movements. (See O'Donnell and Eggemeier [1986] for a review of the various techniques and Hallett [1986] for a thorough review of eye movement research.) While each of these techniques can serve a useful function not all are useful in an applied context. Accordingly, we restrict the discussion to the oculometer (Merchant, Morrissette, and Porterfield,1974) technique which has been used most frequently and successfully in applied situations.

Current eye movement technology permits the investigator to monitor movement of the eyes and, with appropriate calibration, determine where the eyes are pointing, i.e., what instrument was looked at. Additionally, by collecting such data over time, the investigator can determine the scan pattern across an instrument display. An assumption normally made is that dwell time (the length of time the eye stays on an instrument during a look) serves as an index of visual workload: The longer the dwell time, the more difficult to read the instrument. Unfortunately, this is an inferential technique since dwell time could also vary due to the relative importance of the instrument.

WHEN TO USE: Anytime visual/mental workload needs to be analyzed. For cockpit displays and instrumentation in any type of vehicle. (Not suitable in totally free environment although possibly a new head mounted device could offer this.)

SENSITIVITY: Sensitivity appears to be HIGH. Research shows changes in dwell times with a substitution of one instrument for another in a display (Harris & Glover, 1984) or a change in mode of flying (autopilot or manual) (Dick, 1980).

DIAGNOSTICITY: Diagnosticity is MIXED. The technique measures in the visual domain, but manipulations outside vision may increase general workload which can have an effect on dwell times. May require some care in recording the appropriate measures to enable making the necessary inferences (Galanter & Hochberg, 1983).

INTRUSION: Generally LOW. With the oculometer, nothing touches the operator and the operator can move his head normally within a cubic foot without affecting eye movement measures.

IMPLEMENTATION:

Data collection: Since the oculometer is computer driven, data collection is on-line and is straight forward. While the oculometer requires just two channels (for x and y coordinates) multiple channels can be recorded for other parameters such as system state variables, control inputs, etc. (NASA Langley has recorded up to 30 channels simultaneously).

Operator training: None required.

OPERATOR ACCEPTANCE: Little formal work done on this issue. Informal observations have shown airline pilots to have high acceptance. Indeed, training procedures appear to work well based on oculometer recordings.

SAFETY: Generally speaking, safety is HIGH. The operator is unencumbered by any devices. There may be some drying of the corneal surface over long periods of time (3-4 hours) due to the infra-red used in the oculometer.

PHYSICAL SPACE REQUIRED: An area of about 4 x 8 inches is required to mount the infrared light source and the TV camera. Also, space behind the mounting surface is required for the body of the camera and the light source. The lab computer and the recording devices can be located remotely.

PORTABILITY: In normal form currently in use, it is not possible to use as a portable device, although recently developed, head mounted equipment could increase portability. The oculometer has been used in flight as well as fixed and motion based simulators.

INTEGRATION INTO SYSTEMS: Requires technician

RESTRICTIONS: A number of procedures exist for measuring eye movements. While each of these can serve a useful function not all are useful in an applied context.

OTHER: Offers useful training aid. (Jones et al., 1983a; 1983b)

RELATIVE COST OF USE: HIGH. While the recording sessions may be similar to other techniques, the data analysis can be extensive since so many observations are collected.

COMMENTS:

FUTURE DEVELOPMENTS: Addition of pupil diameter measurements would increase the functionality by providing enhanced diagnosticity. The pupil diameter measure and associated technology available with some oculometer units has not matured to the point of being useful.

REFERENCES:

Hallett, P. E. (1986). Eye movements. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance. Vol. I. Sensory processes and perception.* New York: Wiley.

Harris, R. L., Sr., Glover, B. J., & Spady, A. A., Jr. (1986). *Analytic techniques of pilot scanning and their application.* TP-2525. Washington, DC: NASA.

Mazurczak, J., & Pillalamarri, R. S. (1985). *The human engineering eye movement measurement research facility* (TM 6-85) Aberdeen, MD: U. S. Army Human Engineering Laboratory, Aberdeen Proving Ground.

Merchant, J., Morrissette, R., & Porterfield, J. L. (1974). Remote measurement of eye direction allowing subject motion over one cubic foot of space. *IEEE Transactions of Biomedical Engineering, 21,* 309-317.

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance. Vol. II. Cognitive processes and performance.* New York: Wiley.

Heart Rate

DESCRIPTION: Heart rate is related to the amount of physical activity (oxygen requirements), respiration, thermal regulation, and as a result of these activities, heart rate may be related to mental workload. Two measures have been used frequently: heart rate and heart rate variability, the variability measure is discussed separately. Absolute (or mean) heart rate, has not proven to be a direct indication of workload because of the relation of heart rate to a variety of psychological variables which are not normally factored out. Heart rate is sensitive to survival, embarrassment and similar extreme operator emotions, and therefore, closely related to the operator's sense of well-being. As such it is an effective specialized measure. Roscoe and Grieve (1986) and Wierwille and Connor (1983) have independently shown that the measure is sensitive to high stress/ workload in which survival, embarrassment or similar emotions play a role. Unless these latter strong emotions are present, heart rate may not covary with workload (Hart, 1986a).

SENSITIVITY: LOW. There is some controversy with the general measure of heart rate and heart rate variability (O'Donnell & Eggemeier, 1986; Wierwille, 1979); not all investigators have found consistent results, or even results in the same direction. Since heart rate also increases with physical activity, one must take care when measuring mental workload that the measure is not contaminated by high physical activity conditions.

DIAGNOSTICITY: Intrinsically LOW, but can be improved by linking with other information such as the timing of changes in information, response activity, etc.

INTRUSION: Relatively LOW for straight heart rate measures and frequency analysis measures.

IMPLEMENTATION REQUIREMENTS:

Data Collection: Surface electrodes can be used to record the EKG and these may be directly connected by wires to transducers and recording devices or a radio transmitter can be used for more mobile operations.

Heart rate is easily and reliably measured using a peripheral pulse plethysmograph with a sensor attached to the antihelix of the ear.

Operator training: None required.

OPERATOR ACCEPTANCE: Some operators may not be comfortable if electrodes are used.

SAFETY: Any attachment of electrical devices to a human must be accompanied by careful consideration of safety aspects, such as grounding of electrical current, etc.

PHYSICAL SPACE REQUIRED: Little.

PORTABILITY: With a portable transducer and a transmitter, the recording device can be at some distance from the subject.

INTEGRATION INTO SYSTEMS: To improve diagnosticity, it is recommended that system changes and performance be recorded simultaneously.

RELATIVE COST OF USE:

Testing time: Variable.

Equipment: Surface electrodes are required and these may be directly connected by wires to transducers and recording devices or a radio transmitter can be used for more mobile operations.

Set-up and Support: Moderate.

Data Analysis: Can be analyzed simply as beats per minute or more sophisticated analyses can be performed. Graphical representations of heart rate over time might be useful.

COMMENTS:

REFERENCES:

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance. Vol. II. Cognitive processes and performance.* New York: Wiley.

Wierwille, W. W. (1979). Physiological measures of aircrew mental workload. *Human Factors*, *21*, 575-593.

Heart Rate Variability

DESCRIPTION: Mean heart rate is one measure applied to heart rate data. The other measure, heart rate variability (spectral analysis), provides a method to separate out several frequency components and seems to show promise as a measure. One peak, found at 0.35 Hz, represents respiration and a second peak reported at 0.20 Hz represents heart activity related to thermal aspects. For our purposes the important peak, found at 0.1 Hz, seems to be correlated to workload (Sayers, 1973).

SENSITIVITY: Unknown. There is some controversy with the heart rate variability measure (O'Donnell & Eggemeier, 1986; Wierwille, 1979); not all investigators have found consistent results, or even results in the same direction. Some or all of the inconsistency may be due to quite different analysis techniques; Kalsbeek (1973, cited by O'Donnell & Eggemeier, 1986) has reported more than 30 techniques have been used to determine variability. The spectral analysis of heart rate variability seems to show promise since it can separate the various components found in heart rate variance.

DIAGNOSTICITY: Intrinsically LOW, but can be improved by linking with other information such as the timing of changes in information, response activity, etc.

INTRUSION: Relatively LOW for straight heart rate measures and frequency analysis measures.

IMPLEMENTATION REQUIREMENTS:

Data Collection: Surface electrodes are required and these may be directly connected by wires to transducers and recording devices or a radio transmitter can be used for more mobile operations. Speaking will have a disturbing influence on blood pressure and heart rate variability.

Operator training: None required.

OPERATOR ACCEPTANCE: Some operators may not be comfortable if electrodes are used.

SAFETY: Any attachment of electrical devices to a human must be accompanied by careful consideration of safety aspects, such as grounding of electrical current, etc.

PHYSICAL SPACE REQUIRED: Little.

PORTABILITY: With portable transducer and a transmitter, the recording device can be at some distance from the subject.

INTEGRATION INTO SYSTEMS: To improve diagnosticity, it is recommended that system changes and performance be recorded simultaneously.

RELATIVE COST OF USE:

Testing time: Variable.

Equipment: Surface electrodes are required and these may be directly connected by wires to transducers and recording devices or a radio transmitter can be used for more mobile operations.

Set-up and Support: Moderate.

Data Analysis: Fairly sophisticated analysis is required.

REFERENCES:

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance. Vol. II. Cognitive processes and performance.* New York: Wiley.

Wierwille, W. W. (1979). Physiological measures of aircrew mental workload. *Human Factors*, 21, 575-593.

HOS

DESCRIPTION: The Human Operator Simulator (HOS) is a software package for simulation of systems and the humans who operate them. The user builds detailed descriptions of the behavior of the system, the human operator, and any other agents in a simulation scenario and then executes the simulation to produce a detailed timeline of simulated events. Human tasks are defined in a task analytic language which uses basic human performance micro-models to simulate basic cognitive, perceptual, and motor actions of the human. Workload measures may be derived from the timeline (e.g., by determining whether or not all tasks are completed in the available time) or by other user-defined measures involving event data. The first complete version of HOS (HOS-III) was developed by the Navy for use on CDC mainframe computers and is no longer maintained and is not generally available. The only currently available version of HOS (HOS-IV) was developed by the Army Research Institute for use on micro-computers.

SENSITIVITY: HOS simulations can be used to evaluate the effects of any factors and variables that can be explicitly simulated. HOS can easily be made sensitive to issues such as allocation of functions to human or machine, physical placement of control and display devices, and variations of task procedures. It is difficult to achieve sensitivity to individual differences between human operators and to display and information features which are not addressed by the HOS micro-models.

DIAGNOSTICITY: Any observed effects in performance can be traced back to specific conditions represented in the simulation. The development of timing effects can be observed in the detailed event timeline. Simulations can be repeated with variations in key parameters to conduct sensitivity analyses. User-supplied diagnostic software can also be inserted directly into the simulation software.

IMPLEMENTATION REQUIREMENTS:

Data Collection: Development of a complete simulation of a complex system is a substantial effort. It would probably take a new user about a month to perform a simulation study of a fairly simple system with HOS-IV and several months to evaluate a more complex system. Much time and effort can be saved, however, if the new simulation can be adapted from an existing simulation of a similar system.

RELATIVE COST OF USE:

Testing time: It is estimated that typical simulation evaluations using HOS-IV will require from 1 to 6 person months for development of a complex operator simulation, with simpler simulations having a development time of a few days to a few weeks. The computer time required to execute HOS simulations is also a significant factor, with system simulations which are operating at or about the micro-second level, requiring several hours to run.

Equipment: IBM PC-AT and compatibles with 10 megabytes of free disk space, EGA color graphics and a (Microsoft compatible) mouse.

Setup and support: Installation procedure and run-time files provided. Microsoft C version 4.0 or above compiler required.

Data analysis and timelines:

- Task (Rule) analysis gives total time each task executed, the mean time for each task execution, the standard deviation, and the percentage of the simulation which was spent in that task.
- Procedures (Action) timeline, shows, in time order the procedures (actions) that executed during the simulation.
- Event timeline shows the time at which discrete events occurred.

- Object timeline shows that time, object name, attribute name, old and new values for all changes to object attribute values.
- Rule (task) timeline show the time, task status, task type, task number, and task name of all tasks that executed, and the time they ended execution.
- Full timeline combines all of the above timelines. Shows information in the following order: discrete events, rules(tasks), actions (procedures), object attribute changes, within time.

FUTURE DEVELOPMENTS: A new version of HOS (HOS-V) is being developed by the Army Research Institute for use with the HARDMAN-III tools, with availability expected some time in 1992 for a beta test version. HOS-V is being developed specifically for use with the HARDMAN-III tools known as MAN-SEVAL and PER-SEVAL, which are used respectively for evaluating manpower and personnel aspects of system designs. HOS-V will employ a user interface that is consistent with the other HARDMAN-III tools and which, accordingly, has been devised to make it as easy as possible to use. It will incorporate changes to the task description language and micro-model organization in order to improve usability and expandability. It will also include library management functions to facilitate building new simulations out of pieces from prior simulations.

REFERENCES:

Harris, R., Iavecchia, H., Ross, L., and Shaffer, S. "Microcomputer Human Operator Simulator (HOS-IV)," Proceedings of the 1987 Human Factors Society Meeting, Santa Monica, CA: Human Factors Society, 1987.

Glenn, F. "The Human Operator Model Manager for System Design" Proceedings of Human Factors Society 32nd Annual Meeting, Santa Monica, CA: Human Factors Society, 1988.

Harris, R., Iavecchia, H., Zaklad, A., and Glenn, F. Human Operator Simulator". In Karwowski, W. (ed.), Trends in Ergonomics/Human Factors III, New York: North-Holland. 1986.

Glenn, F., Zaklad, A., and Wherry, R. "Human Operator Simulation in the Cognitive Domain" Proceedings of the Human Factors Society 27th Annual Meeting, Santa Monica, CA: Human Factors Society, 1984.

Lane, N., Strieb, M., Glenn, F., and Wherry, R. "The Human Operator Simulator: An Overview" In Manned Systems Design: Methods, Equipment, and Applications, J. Moraal and K.-F. Kraiss (Eds.). New York: Plenum Press. 1981.

AVAILABILITY:

U.S. Army Research Institute for the Behavioral and Social Sciences
Systems Research Laboratory
5001 Eisenhower Ave.
Alexandria, VA 22333-5600

- Users' and programmers' guides available for HOS-IV
- Software requires PC-AT type computer with C compiler

McCracken-Aldrich

DESCRIPTION: Special-purpose methodology for evaluating crew overload in an aviation setting. In a top-down approach, a mission is broken down into segments, segments into functions, and functions into performance elements (comparable to a task level). For each performance element, subject matter experts assign workload ratings (on a scale of 1 to 7) for cognitive, visual, auditory, and psychomotor channels as well as performance duration and the associated subsystem. A scenario timeline is generated using segment decision rules that define what tasks are being performed in what sequence for three primary ongoing activities including flight control, flight support, and mission-related activities.

SENSITIVITY: Crew workload at the task level.

DIAGNOSTICITY: Determine how workload varies as a function of crew size. Identify subsystems associated with high workload.

INTRUSION:

IMPLEMENTATION REQUIREMENTS:

Detailed task analysis for a given mission.

Workload ratings for cognitive, visual, auditory, and psychomotor channels for each operator function as well as task time and associated subsystem.

OPERATOR ACCEPTANCE:

RELATIVE COST OF USE:

Testing time: several months to develop

Equipment: Perkin-Elmer

Setup and support: FORTRAN programmer.

Data analysis:

Tools:

COMMENTS: Methodology for assessing workload in early system design stages. Applied to the LHX to evaluate the effects of one versus two crewmembers on system performance. See TIS on TAWL.

FUTURE DEVELOPMENTS: See the TAWL TIS.

REFERENCES:

McCracken, J. H., & Aldrich, T. B. (1984). Analysis of selected LHX mission functions: Implications for operator workload and system automation goals (TNA ASI479-24-84). Fort Rucker, AL: Anacapa Sciences, Inc.

AVAILABILITY:

Chief
Army Research Institute
Aviation Research and Development Activity
ATTN: PERI--IR (Mr. Charles A. Gainer)
Fort Rucker, AL 36362-5354

MicroSaint

DESCRIPTION: MicroSAINT is a task network model in which activities are represented in a diagram as nodes and the arrows between nodes show the sequence in which those activities are performed. The task is the basic building block of a MicroSAINT model. A single model may have up to 400 tasks, each of which is entered into the model by entering data after selecting the Modify Task option on the main menu. MicroSAINT in itself contains no predefined workload model but through the use of its scripting language, functions can be written to calculate workload estimates at any time during the simulation. These workload estimates are simply the mathematical manipulations of the workload estimates obtained from applying a specific workload assessment methodology a(e.g., SWAT) to individual tasks. Operator workload estimates are then incorporated into the simulation by assigning the individual task values to the task beginning effect. When a task executes, the values in the task beginning effect become active and can be accessed by the user-written functions.

SENSITIVITY: The results obtained from a MicroSAINT simulation are only as sensitive as the operator workload estimates which are incorporated into the task network model.

DIAGNOSTICITY: Again, this depends mainly on the operator workload estimation methodology. However, once the simulation has been created it can be easily modified to predict the effects of varying the sequence, structure, or duration of individual tasks in the network.

INTRUSION: N/A

IMPLEMENTATION REQUIREMENTS:

Data Collection: Before the task workload estimation process is exercised, a thorough task analysis must be performed to provide the basis of the MicroSAINT task network model. The translation of the task analysis results to MicroSAINT is a fairly straightforward process; however, the embellishment of the model with user defined functions can become very complex.

Operator Training: N/A

RELATIVE COST OF USE:

Testing time: It is estimated that a moderately complex model containing a minimal workload function would require a development time of two person weeks.

Equipment: IBM PC

Setup and support: MicroSAINT is delivered with an installation program which allows the system to be setup in less than an hour. Technical support is available for Micro Analysis and Design.

Data analysis: MicroSAINT provides minimal statistical analysis and graphing functions, but MicroSAINT output files (tab delimited) are easily read by other statistical software packages.

Tools: Analysis of simulation results sometimes require the use of other software packages.

FUTURE DEVELOPMENTS: A new variant of MicroSAINT, the Workload Assessment Aid (WAA), is currently being developed by the Army explicitly for evaluating operator workload. A Macintosh-based version is also under development

REFERENCES:

Dahl, S. G., Drews, C. W., Kelly, K. J., & Plott, C. C. (1987). Micro SAINT: A simulation tool for the human factors professional. *CSTG Bulletin, 14,* 14-17.

Laughery, K. R., Jr., Drews, C., Archer, R., & Kramme, K. (1986). A MicroSAINT simulation analyzing operator workload in a future attack helicopter. In *National Aerospace and Electronics Conference* (pp. 896-903). Dayton, OH: IEEE.

Laughery, K. R. (1989). Task network modeling as a.basis for analyzing operator workload In *Proceedings of the Human Factors Society 33rd Annual Meeting*. Santa Monica, CA: Human Factors Society.

AVAILABILITY:

Micro Analysis and Design, Inc.
9132 Thunder Head Dr.
Boulder, Co 80302

- Runs on IBM PC and compatibles
- Diskettes and manuals

Modified Cooper Harper (MCH)

DESCRIPTION: The MCH is used to obtain ratings from 1-100 via a decision tree structure. Although derived from the Cooper-Harper, it was designed to be applicable to a broad number of operational environments (i.e., it is not specifically a pilot rating scale). It can be used in real-time operation.

SENSITIVITY: The scale has been reported to be sensitive to differences in task loading.

DIAGNOSTICITY: The MCH gives a global rating of workload.

INTRUSION: Little, although it does require a judgment. There was concern (as with most subjective measures) that the judgment might interfere with flight duties, but ratings were able to be obtained real-time.

IMPLEMENTATION REQUIREMENTS:

Data Collection: Some method for collecting the ratings is needed -- either a 10 key pad or communications medium with which the operator can report the rating verbally. A copy of the scale for reference is also useful.

Operator Training: The operators must be given an opportunity to become familiar with the rating scale, therefore some practice is necessary, although the scale is reported to be easy to understand.

OPERATOR ACCEPTANCE: The scale has been reported to be well received by experimental subjects who were pilots.

SAFETY: Plans must be made as to what to do if the operator is too busy to give a rating. Ratings should be secondary to the primary concern with operational safety (e.g., flying a plane or controlling a land vehicle).

PHYSICAL SPACE REQUIRED: n/a

PORTABILITY: n/a

INTEGRATION INTO SYSTEM: n/a

RESTRICTIONS: n/a

RELATIVE COST OF USE:

Testing time: Minimal.

Equipment: Minimal.

Setup and support: Minimal.

Data analysis: Descriptive and inferential statistics can be used. Graphical representations are useful. Caution is advised in assuming an interval scale, therefore non-parametric analysis may be more appropriate.

COMMENTS:

REFERENCES:

Wierwille, W. W., & Casali, J. G. (1983). A validated rating scale for global mental workload measurement application. In *Proceedings of the Human Factors Society 27th Annual Meeting* (pp. 129-133). Santa Monica, CA: Human Factors Society.

Wierwille, W. W., Casali, J.G., Connor, S. A., & Rahimi, M. (1985). Evaluation of the sensitivity and intrusion of mental workload estimation techniques. In W. Roner (Ed.), *Advances in Man-Machine Systems Research. Vol. 2,* pp. 51-127). Greenwich, CT: J.A.I. Press.

Wierwille, W. W., Skipper, J., & Reiger, C. (1984). Decision tree rating scales for workload estimation. Theme and variations (NASA-CP-2341). In *Proceedings of the 20th Annual Conference on Manual Control* (pp. 73-84). Washington, D.C: NASA.

Open-Ended Questionnaires

DESCRIPTION: Questionnaires are forms in which written questions are asked in a fixed order and format and to which respondents write their answers. The questions may be open-ended, allowing respondents to write in their own words and make any answer, or close-ended, where the choice of answers has been previously established, such as multiple choice or true and false. Questionnaires should be used whenever possible to obtain the subtle, detailed information that might not be obtained from rating scales.

SENSITIVITY: Variable.

DIAGNOSTICITY: Detailed information that might not otherwise be obtained can be drawn from interviews.

INTRUSION: Depends on the length of the questionnaire, but for the most part, questionnaires will not be appropriate during real-time operation. Answering questions will require attention and will distract the operator from the primary task.

IMPLEMENTATION REQUIREMENTS:

Data Collection: The most common implementation is via paper and pencil, however, questionnaires can be administered via computer if available.

Operator Training: Minimal.

OPERATOR ACCEPTANCE: In general, questionnaires are well-accepted. However, if questions are not clear, or operators are asked too many questions too often, acceptance may decrease.

SAFETY: Plans must be made as to what to do if the operator is too busy to give a rating. Ratings should be secondary to the primary concern with operational safety (e.g., flying a plane or controlling a land vehicle).

PHYSICAL SPACE REQUIRED: n/a

PORTABILITY: n/a

INTEGRATION INTO SYSTEM: n/a

RESTRICTIONS: n/a

RELATIVE COST OF USE:

Testing time: Can vary in time, depending on how many questions are asked and whether they are open or close-ended.

Equipment: Paper and pencil (or can be computer-based).

Setup and support: Careful development is needed. Little support.

Data analysis: Qualitative and quantitative.

COMMENTS:

REFERENCES:

Dyer, R., Matthews, J., Wright, C., & Yudowitch, K. (1976). *Questionnaire Construction Manual* (TCATA DAHC-19-74-C-0032). Ft. Hood, TX: ARI.

U.S. Army Test and Evaluation Command (1976). *Questionnaire and interview design, Subjective testing techniques* (TECOM Pam 602-1, Vol. I). Aberdeen Proving Ground: USATECOM.

Meister, D. (1985). *Behavioral analysis and measurement methods.* New York: John Wiley and Sons.

OW and Prospective OW

DESCRIPTION: The overall workload (OW) scale is a unidimensional bipolar rating scale which an operator can use to give an absolute estimate of the workload experienced during a particular mission segment. The scale consists of a horizontal line divided into 20 equal intervals; the words "low" and "high" are placed, respectively, at the left and right ends of the scale. Numerical values, assigned by the analyst, range form 0 to 100.

SENSITIVITY: The scale has been shown to be sensitive to differences in task loading for a variety of different tasks, systems, and operational environments.

DIAGNOSTICITY: OW gives only a global indication of the overall workload experienced by the operator.

INTRUSION: Little, though it requires that the operator give an absolute judgment. Even so, studies have shown that OW ratings can be obtained in real time without interfering with the operator's performance.

IMPLEMENTATION REQUIREMENTS:

Data Collection: The OW scale can be administered during (real time), after (retrospectively), or before (prospectively) the operator performs the task of interest. The operator ratings can be obtained verbally, by paper and pencil, or electronically via a keypad.

Operator Training: Some practice in using the scale and understanding the operational meaning of the scale (and of the concept of workload) is helpful.

OPERATOR ACCEPTANCE: High

SAFETY: Plans must be made as to what to do if the operator is too busy to give a real-time rating. Normally, the analyst can ask for a retrospective rating at some period of time after the task of interest has been completed.

PHYSICAL SPACE REQUIRED: n/a

PORTABILITY: n/a

INTEGRATION INTO SYSTEM: n/a

RESTRICTIONS: n/a

RELATIVE COST OF USE:

Testing time: Minimal.

Equipment: Minimal.

Setup and support: Minimal.

Data analysis: Minimal.

COMMENTS: When used retrospectively, after a long delay, the operator should be aided in recreating the experiences associated with the task when it was previously performed; audio and video recordings of task performance are helpful in this regard. When used prospectively, the operator or subject matter expert should be aided in creating a useful representation of the task as well as the system and operating environment which form the context of the task that is to be rated. In this latter case, the ratings of workload are made to descriptions of tasks and events that have not yet been personally experienced by the individual making the ratings (see Eggleston & Quinn, 1984).

REFERENCES:

Byers, J.C., Bittner, A.C., Jr., Hill, S.G., Zaklad, A.L., & Christ, R.E. (1988). Workload assessment of a remotely piloted vehicle (RPV) system. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 1145-1149). Santa Monica, CA: Human Factors Society.

Eggleston, R.G., & Quinn, T.J. (1984). A preliminary evaluation of a projective workload assessment procedure. In *Proceedings of the Human Factors Society 28nd Annual Meeting* (pp. 695-699). Santa Monica, CA: Human Factors Society.

Hill, S.G., Zaklad, A.L., Bittner, A.C., Jr., Byers, J.C., & Christ, R.E. (1988). Workload assessment of a mobile air defense missile system. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 1068-1072). Santa Monica, CA: Human Factors Society.

Iavecchia, H.P., Linton, P.M., & Byers, J.C. (1989). Workload assessment during day and night missions in a UH-60 Blackhawk helicopter simulator. In *Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 1481-1485). Santa Monica, CA: Human Factors Society.

Vidulich, M.A., & Tsang, P.S. (1987). Absolute magnitude estimation and relative judgement approaches to subjective workload assessment. In *Proceedings of the Human Factors Society 31st Annual Meeting* (pp. 1057-1061). Santa Monica, CA: Human Factors Society.

AVAILABILITY: The OW scale is one of the subscales used during the construction of the NASA-TLX. See the TLX TIS for more information.

Pupil Diameter

DESCRIPTION: It is well known that pupil diameter varies with a number of physiological and psychological variables. The iris of the eye changes diameter as a function of mental and physical states. On the one hand, pupil diameter is one of the adaptive mechanisms of the eye to control the amount of light entering the eye. Depending on the ambient light level, pupil diameter will vary from about 1 mm upwards to 8 mm in total darkness. While not the primary adaptive mechanism, pupil diameter does serve as an important role for depth of field, much like the shutter of a camera. On the other hand and of more interest in the present discussion, pupil diameter also varies as a result of psychological variables. Pupil diameter is controlled by smooth muscles which are driven through the autonomic nervous system; the autonomic system, in turn, is influenced by cortical activity (Moses, 1970). Thus, the known physiology is consistent with use of pupil diameter as a measure of mental workload and momentary mental states. Additionally, because the pupil varies with light levels independently of workload states, one must be careful to keep ambient light at a constant to avoid contamination of the data. While it should be possible to remove this effect analytically, this has not been done.

SENSITIVITY: HIGH. Under controlled conditions, the measure has been shown to be sensitive to difficulty of auditory tasks.

DIAGNOSTICITY: Like many techniques, diagnosticity is mixed and depends upon the ingenuity of the investigator. Combined with multiple measures, pupil diameter measures show capability of fine grain, detailed Diagnosticity.

INTRUSION: Intrusion is LOW.

IMPLEMENTATION REQUIREMENTS:

Data Collection: Data collection can be computerized which simplifies the data acquisition step.

Operator training: None required.

OPERATOR ACCEPTANCE: Operator acceptance should be high since the measure is unobtrusive, but little formal work has been reported.

SAFETY: Safety is HIGH since nothing is required to touch the subject.

PHYSICAL SPACE REQUIRED: Some space needed for equipment.

PORTABILITY: Normally, the measurement has been used as a laboratory technique.

INTEGRATION INTO SYSTEMS:

RESTRICTIONS: At present, the restrictions are in terms of adjusting for eye movements. When viewed straight on, the pupil appears circular. As the eye moves horizontally, this circle becomes an ellipse with the long axis oriented vertically and the short axis of horizontal orientation. Vertical eye movements will produce a similar ellipse but with a different orientation of the long and short axes. If eye movements were only in the horizontal or vertical directions, adjustment would be easy; simply measure the long axis to determine the diameter. Oblique movements, however, require knowledge of eye position. There does not appear to be a commercial apparatus which does both pupil diameter and eye position.

RELATIVE COST OF USE:

Testing time: Since the data are collected on line with the task, the measure does not affect the actual testing time, although the testing time may be affected by calibration measures.

Equipment: The typical technique has been to use video tape and do the analysis later. Commercially available devices can be used which provide on-line analysis and or electronic recording suitable for computer analysis.

Setup and Support: Modest.

Data Analysis: Data analysis techniques are similar to evoked potential with a number of trials averaged using a timing mark to begin the averaging.

COMMENTS:

FUTURE DEVELOPMENTS:

If the eye movement correction analysis can be accomplished and the pupil reflex (ambient light levels) effect can be removed analytically, then the technique could be of considerable value in an applied context.

REFERENCES:

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91*, 276-292.

Moses, R. A. (1970). *Adler's Physiology of the eye. Clinical application.* St. Louis: C. V. Mosby

SIMWAM

DESCRIPTION: The Simulation for Workload Assessment and Manning (SIMWAM) methodology (Kirkpatrick, Malone & Andrews, 1984) is based on both MicroSaint (see MicroSaint TIS) and WAM) see Tr/Ta Task Analysis TIS). However, it has been specifically developed to make it especially suitable for examining manpower issues, as well as individual operator workload, in complex multi-operator systems, where interactions among crewmembers is a critical feature of system operations.

SENSITIVITY: Varies.

DIAGNOSTICITY: Varies.

INTRUSION: n/a

IMPLEMENTATION REQUIREMENTS:

Data Collection: Requires operational sequence diagrams and detailed task analysis with task performance times.

Operator training: n/a

OPERATOR ACCEPTANCE: n/a

SAFETY: n/a

PHYSICAL SPACE REQUIRED: n/a

PORTABILITY: n/a

INTEGRATION INTO SYSTEMS: n/a

RESTRICTIONS: n/a

RELATIVE COST OF USE:

Testing time: n/a

Equipment: n/a

Setup and Support:

Data Analysis:

COMMENTS: SIMWAM has been used to assess workload and manpower issues for an aircraft carrier's aircraft operations management system (Malone, Kirkpatrick & Kopp, 1986). The interactive nature of SIMWAM allows the analyst to evaluate alternative system design or modification concepts involving manpower reduction, cross-training, automation, task modification, or function allocation.

REFERENCES:

Kirkpatrick, M., Malone, T.B., & Andrews, P.M. (1984). Development of an interactive microprocessor-based workload evaluation model (SIMWAM). In *Proceedings of the Human Factors Society 28th Annual Meeting* (pp. 78-80). Santa Monica, CA: Human Factors Society.

Malone, T.B., Kirkpatrick, M., & Andrews, P.M. (1986). Human factors engineering impact of system workload and manning levels.. In *Proceedings of the Human Factors Society 30th Annual Meeting* (pp. 763-767). Santa Monica, CA: Human Factors Society.

Sternberg Memory Task

DESCRIPTION: Sternberg Memory task has been used as a secondary task to reflect OWL levels in primary tasks. The subject is presented with a set of digits or letters to memorize. Subsequently, the subject is presented with a test digit or letter and must judge whether this digit or letter was contained in the previous memorized set.

SENSITIVITY: Has been shown to be sensitive to differences in aircraft task difficulty as defined by different flight maneuvers (e.g., holding pattern versus approach pattern). It is generally seen as a secondary task which reflects cognitive central processing loads.

DIAGNOSTICITY: Gives a global measure of workload. However, it can be used in controlled laboratory type situations to distinguish between perceptual/central processing task loading and response task loading for a primary task.

INTRUSION: Little, although there may be situations where the Sternberg Memory task might interfere with the primary task. Studies have used the task in flight simulators without any significant interference with flight duties.

IMPLEMENTATION REQUIREMENTS:

Data Collection: Some method is needed to collect the operator's responses and store the data. It is preferable to store the data with some type of time stamp in order to reference the operator's responses to specific events.

Operator Training: Operators must be given an opportunity to practice the Sternberg Memory task as well as establish a baseline for their responses to be used for data analysis.

OPERATOR ACCEPTANCE: Pilots in flight simulators have been receptive to performing the Sternberg Memory task within flight scenarios.

SAFETY: Plans must be made as to what to do if operators are too busy to respond to probe letters or digits. The task should be secondary to the primary concern with operational safety (e.g., flying an operational aircraft or controlling a land vehicle).

PHYSICAL SPACE REQUIRED: n/a

PORTABILITY: Varies as a function of the particular set-up (i.e., hardware and software) for the Sternberg Memory task.

INTEGRATION INTO SYSTEM: Varies as a function of the particular set-up (i.e., hardware and software) for the Sternberg Memory task.

RESTRICTIONS: Variable - see Safety

RELATIVE COST OF USE:

Testing time: A sufficient number of responses is needed to nave reliable data for each operator within a test session. It is advisable to change the memory set after 20-30 trials to a different one in order to maintain sensitivity for OWL levels.

Equipment: Minimal - medium.

Setup and support: Minimal - medium.

Data analysis: Descriptive and inferential statistics can be used for reaction time scores for positive and negative probe items. Graphical representations are useful.

COMMENTS:

REFERENCES:

Wickens, C. D., Hyman, F., Dellinger, J, Taylor, H., & Meador, M. (1986). The Sternberg memory search task as an index of pilot workload. *Ergonomics, 29*, 1371-1383.

Structured Interviews

DESCRIPTION: Structured interviews are those in which the questions asked are carefully planned (structured) beforehand and followed in the interview process. These interviews are held with the operators to obtain the subtle, detailed information that might not be obtained from rating scales. The questions for structured interviews are available on paper, so that different interviewers can ask the same questions.

SENSITIVITY: Variable.

DIAGNOSTICITY: Detailed information that might not otherwise be obtained can be drawn from interviews.

INTRUSION: Most appropriate to interview after the test session or operational sequence to obtain the most information and reflections of the operator.

IMPLEMENTATION REQUIREMENTS:

Data Collection: An interviewer who is familiar with the system/operation under study is necessary so that subtle information regarding OWL can be drawn from the interviewee. Video or audio recording, or paper and pencil transcription.

Operator Training: Minimal, although operators may become more comfortable with expressing opinions as they become more experienced in being interviewed.

OPERATOR ACCEPTANCE: Interviewing is a well established method of obtaining subjective information. Cooperation can generally be expected to be good, but there may be individuals who are uncomfortable and will be uncooperative. In addition, there can be the problem of individuals who are unwilling to give negative reports. Practitioners should be aware of possibly misleading information obtained through interviews.

RELATIVE COST OF USE: The highest cost is the time invested by the interviewer.

Testing time: Variable, although less than 30 minutes is recommended.

Equipment: Recording equipment, if used.

Setup and support: Minimal.

Data analysis: Qualitative.

COMMENTS:

REFERENCES:

Dyer, R., Matthews, J., Wright, C., & Yudowitch, K. (1976). *Questionnaire Construction Manual* (TCATA DAHC-19-74-C-0032). Ft. Hood, TX: ARI.

U.S. Army Test and Evaluation Command (1976). *Questionnaire and interview design, Subjective testing techniques* (TECOM Pam 602-1, Vol. I). Aberdeen Proving Ground: USATECOM.

Meister, D. (1985). *Behavioral analysis and measurement methods.* New York: John Wiley and Sons.

Ericsson, K. A. (1984). *Protocol Analysis.* Cambridge, MA: MIT Press.

SWAT and Prospective SWAT

DESCRIPTION: SWAT uses the three dimensions of time load, mental effort load, and psychological stress load to assess workload. For each dimension, there are three operationally defined levels. SWAT has two parts: 1) a card sort procedure where the operator determines the rank order of all combinations of the three levels of the three dimensions; and 2) an event scoring part where the operator makes ratings of the three dimensions. Conjoint analysis is used to obtain a global workload rating between 0 and 100.

SENSITIVITY: SWAT has been demonstrated to be sensitive to task loading in a number of different types of tasks.

DIAGNOSTICITY: SWAT gives a global rating of workload. However, the three subscales can be examined individually and used for diagnostic purposes.

INTRUSION: Little, although it does require a judgment. There was concern (as with most subjective measures) that the judgment might interfere with flight duties, but ratings were able to be obtained real-time.

IMPLEMENTATION REQUIREMENTS:

Data Collection: The card sort procedure can take up to an hour to perform. The SWAT event ratings can be administered during (real time), after (retrospectively), or before (prospectively) the operator performs the task of interest. The operator ratings can be obtained verbally, by paper and pencil, or electronically via a keypad.

Operator Training: Practice is needed for the operators to become familiar with the operational definitions and the giving of ratings.

OPERATOR ACCEPTANCE: SWAT has been used successfully in aviation and other application. However, cooperation and motivation is the key to obtaining a valid card sort which are the most difficult aspect of this technique.

SAFETY: Plans must be made as to what to do if the operator is too busy to give real-time ratings. Real-time ratings should be secondary to the primary concern with operational safety (e.g., flying a plane or controlling a land vehicle).

PHYSICAL SPACE REQUIRED: n/a

PORTABILITY: n/a

INTEGRATION INTO SYSTEM: n/a

RELATIVE COST OF USE:

Testing time: Card sort can take up to an hour, while the event ratings can be obtained very quickly.

Equipment: Whatever equipment is chosen for data collection. Computer access is necessary for data reduction and analysis.

Setup and support: Careful administration is required, particularly for card sort.

Data analysis: Descriptive and inferential statistics can be used. Parametric statistics are appropriate since conjoint scaling provides an interval scale and they have been used to examine significant differences between mission segments or task variables.

COMMENTS: When used retrospectively, after a long delay, the operator should be aided in recreating the experiences associated with the task when it was previously performed; audio and video recordings of task performance are helpful in this regard. When used prospectively, the operator or subject matter expert should be aided in creating a useful representation of the task as well as the system and operating environment which form the context of the task that is to be rated. In this latter case, the ratings of workload are made to descriptions of tasks and events that have not yet been personally experienced by the individual making the ratings (see Eggleston & Quinn, 1984).

REFERENCES:

Armstrong Aerospace Medical Research Laboratory (1987, June). *Subjective Workload Assessment Technique (SWAT): A User's Guide*. Dayton, OH: AAMRL, Wright Patterson AFB.

Eggleston, R.G., & Quinn, T.J. (1984). A preliminary evaluation of a projective workload assessment procedure. In *Proceedings of the Human Factors Society 28nd Annual Meeting* (pp. 695-699). Santa Monica, CA: Human Factors Society.

Reid, G. B., Eggemeier, F., & Nygren, T. (1982). An individual differences approach to SWAT scale development. In *Proceedings of the Human Factors Society 26th Annual Meeting* (pp. 639-642). Santa Monica, CA: Human Factors Society.

Reid, G. B., Shingledecker, C. A., & Eggemeier, F. T. (1981). Application of conjoint measurement to workload scale development. In *Proceedings of the Human Factors Society 25th Annual Meeting* (pp. 522-525). Santa Monica, CA: Human Factors Society.

AVAILABILITY:

Workload and Ergonomics Branch
Human Engineering Division
Department of the Air Force
Armstrong Aerospace Medical Research Laboratory
Wright-Patterson Air Force Base, Ohio 45433-6573

- Instruction manuals and materials
- Computer program for data analysis, runs on IBM PC and compatibles

TAWL

DESCRIPTION: For a given crewmember and scenario, the Task Analysis/Workload (TAWL; Bierbaum, Fulford, and Hamilton, 1990; Hamilton, Bierbaum, and Fulford, 1991) methodology predicts operator overload using a data base of information produced from a task and workload analysis (see TIS on the predecessor McCracken-Aldrich model). Using a top-down approach, a mission is broken down into phases, phases into segments, segments into functions, and functions into tasks. For example, in an AH-64 evaluation (Szabo & Bierbaum, 1986), seven mission phases, 49 segments, 153 functions, and 653 tasks were identified. For the task analysis, the duration of each task is specified as well as the associated crewmember and subsystem. For the workload analysis, a subject matter expert assigns workload ratings (on a scale from 1 to 7) to the auditory, visual, visual-aided, kinesthetic, cognitive, and psychomotor channels for each task. A scenario is defined using segment and function rules. Segment rules specify what functions will be performed sequentially and concurrently by each crewmember within a specific segment. Similarly, function rules specify what tasks will be performed sequentially and concurrently by each crewmember within a specific function. Randomly-occurring tasks are also defined. A scenario timeline is then generated using the segment and function rules. Independent channel workload is estimated for each time snapshot.

SENSITIVITY: Operator workload at the task level. Can also identify subsystems associated with high workload.

DIAGNOSTICITY: Determine how workload varies across time, crewmembers, channel components (e.g., cognitive, psychomotor), and subsystems.

INPUTS:

Detailed task analysis defining the low-level task activities required for a mission including task times.

Workload ratings for auditory, visual, visual-aided, kinesthetic, cognitive, and psychomotor channels on a scale of 1 to 7 for each low-level task activity.

Scenario decision rules indicating the activities to be performed by each operator.

OUTPUTS:

Generates a timeline of low-level activities and predictions of workload at fixed half-second intervals and summary reports of workload statistics, overloads, subsystem use, and subsystem impact on the workload of up to four crewmembers.

RELATIVE COST OF USE:

Testing time: 6 months to develop a baseline model

Equipment: Perkin-Elmer for original TAWL software; IBM-PC compatible for the microcomputer implementation known as TAWL Operator Simulation System (TOSS; Hamilton, Bierbaum, and Fulford, 1991; Fulford, and Hamilton; and Bierbaum, 1990).

Setup and support:

Data analysis:

Tools:

COMMENTS: TAWL has primarily been applied to predict the impact of system design upgrades on workload in Army aviation settings. Recent applications include various Army ground-based crewstations. Computer implementation of this methodology is necessary. The original TAWL software was developed on a Perkin-Elmer minicomputer. The TAWL Operator Simulation System (TOSS) is a microcomputer implementation of the methodology that employs a menu-driven user-computer interface

(Bierbaum, Fulford, and Hamilton; 1989). MicroSaint can also be used to implement the methodology.

REFERENCES:

Bierbaum, C.R., Fulford, L.A., & Hamilton, D.B. (1990). *Task Analysis/Workload (TAWL) User's Guide - Version 3.0* (Research Product 90-15). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD S221 865)

Fulford, L.A., Hamilton, D.B., & Bierbaum, C. R. (1990). TAWL operator simulation system (TOSS) Version 4.0. *Proceedings of the Human Factors Society, 34th Annual Meeting* (p. 1096). Santa Monica, CA: Human Factors Society.

Hamilton, D.B., Bierbaum, C.R., & Fulford, L.A. (1991). *Task Analysis/Workload (TAWL) User's Guide - Version 4.0* (Technical Report ASI690-330-90). Fort Rucker, AL: Anacapa Sciences, Inc.

Hamilton, D.B., Bierbaum, C. R., & Fulford, L.A. (1991). Task Analysis/Workload (TAWL): A methodology for predicting operator workload. Proceedings of the Human Factors Society, 35th Annual Meeting (pp. 1117-1121). Santa Monica, CA: Human Factors Society.

Szabo, S. M., & Bierbaum, C. R. (1986). A comprehensive task analysis of the AH-64 mission with crew workload estimates and preliminary decision rules for developing an AH-64 workload prediction model. Vol. I. (ASI678-204-86[B]). Ft. Rucker, AL: Anacapa Sciences, Inc.

AVAILABILITY:

Chief
Army Research Institute
Aviation Research and Development Activity
Attn: PERI-IR (Mr. Charles A. Gainer)
Ft. Rucker, AL 36362-5354

Time Estimation Secondary Task

DESCRIPTION: Time estimation has been used as a secondary task to reflect OWL levels with primary tasks. The subject keeps track of time either by generating a specific time interval or by estimating the duration of a time interval at its conclusion. Typically, subjects are required to generate 10 second time intervals (time production procedure). It is assumed under high workload conditions that subjects will underestimate the passage of time as reflected by their responses (i.e., longer time estimates).

SENSITIVITY: Has been shown to be sensitive to differences in aircraft task difficulty as defined by mission scenarios which varied the task demands for particular flight duties, for example, detection of emergency situations.

DIAGNOSTICITY: Gives a global measure of workload. It can be examined with respect to specific instances within a flight scenario to determine the OWL associated with a specific flight task. However, such a determination requires repeated administration of the Time Estimation task in association with the specific flight task to produce reliable results.

INTRUSION: Little, although there may be situations where Time Estimation might interfere with the primary task. Studies have used Time Estimation in flight simulators without any interference with flight duties.

IMPLEMENTATION REQUIREMENTS:

Data Collection: Some method is needed to collect the operator's responses (time production method) and store the data. It is preferable to store the data with some type of time stamp in order to reference the Time Estimation responses to specific events.

Operator Training: Operators must be given an opportunity to practice Time Estimation as well as establish a baseline for their time estimates to be used for data analysis.

OPERATOR ACCEPTANCE: Pilots in flight simulators have been receptive to performing Time Estimations within flight scenarios.

SAFETY: Plans must be made as to what to do if operators are too busy to produce time estimates. The task should be secondary to the primary concern with operational safety (e.g., flying an operational aircraft or controlling a land vehicle).

PHYSICAL SPACE REQUIRED: n/a

PORTABILITY: Varies as a function of the particular set-up (i.e., hardware and software) for the Time Estimation task.

INTEGRATION INTO SYSTEM: Varies as a function of the particular set-up (i.e., hardware and software) for the Time Estimation task.

RESTRICTIONS: Variable - see Safety

RELATIVE COST OF USE:

Testing time: A sufficient number of time estimates is needed to have reliable data for each operator within a test session.

Equipment: Minimal - moderate.

Setup and support: Minimal - moderate.

Data analysis: Descriptive and inferential statistics can be used for time estimate scores and variability scores. Graphical representations are useful.

COMMENTS:

REFERENCES

Hart, S. G. (1978). Subjective time estimation as an index of workload. In *Proceedings of the symposium on man-system interface: Advances in workload study* (pp. 115-131).

Wierwille, W.W., Casali, J. G., Connor, S. A., & Rahimi, M. (1985). Evaluation of the sensitivity and intrusion of mental workload estimation techniques. In W. Roner (Ed.), *Advances in Man-Machine Systems Research* (Volume 2, pp. 51-127). Greenwich, CT: J.A.I. Press.

TLX and Prospective TLX

DESCRIPTION: NASA-TLX is a multidimensional scale that uses an individual weighting procedure to reduce between-subject variability. It was derived from the NASA-Bipolar scales. It is comprised of two procedures: 1) six rating scales covering different dimensions of workload used to rate OWL; and 2) the "Sources of Workload Evaluation" using paired comparisons of the six dimensions to obtain individual weightings of the dimension importance to workload for any task. The ratings and weightings are combined to produce a global workload rating between 0 and 100.

SENSITIVITY: Has been demonstrated to be sensitive to differences in task loading in a number of different types of tasks.

DIAGNOSTICITY: NASA-TLX gives a global rating of workload. However, the six subscales can potentially be examined individually and used for diagnostic purposes.

INTRUSION: Very little, although it does require a judgment. There was concern (as with most subjective measures) that the judgment might interfere with flight duties, but ratings were able to be obtained real-time.

IMPLEMENTATION REQUIREMENTS:

Data Collection: A "Sources of Workload Evaluation" will be obtained for each task under study. The procedure uses only 15 paired comparisons and does not require much time to accomplish. The six TLX scales used to obtain ratings can be administered during (real time), after (retrospectively), or before (prospectively) the operator performs the task of interest. The operator ratings can be obtained verbally, by paper and pencil, or electronically via a keypad.

It has been suggested that an alternative to collecting "Sources of Workload Evaluation" is to use Raw TLX (i.e., non-weighted TLX scores) (Byers, Bittner and Hill, 1989).

Operator Training: Some practice in using and understanding the operational descriptions of the scales would be helpful.

OPERATOR ACCEPTANCE: Has been used successfully in real-time and post-flight aviation applications.

SAFETY: Plans must be made as to what to do if the operator is too busy to give real-time ratings. Real-time ratings should be secondary to the primary concern with operational safety (e.g., flying a plane or controlling a land vehicle).

PHYSICAL SPACE REQUIRED: n/a

PORTABILITY: n/a

INTEGRATION INTO SYSTEM: n/a

RESTRICTIONS: n/a

RELATIVE COST OF USE:

Testing time: The "Sources of Workload Evaluation" takes on the order of 10 minutes to make paired comparisons. The six ratings would not take significant time if the operators were familiar with the scale descriptions.

Equipment: Can be obtained via paper and pencil, or via computer. Video recording equipment is necessary in order to tape operator activity for use in post-test visual recreation.

Setup and support: Minimal.

Data analysis: The weighting and global measure computation can be done by hand, although a computer would be helpful. Descriptive and inferential statistics can be

applied. Parametric and non-parametric statistics have been used to examine significant differences between mission segments or task variables.

COMMENTS: When used retrospectively, after a long delay, the operator should be aided in recreating the experiences associated with the task when it was previously performed; audio and video recordings of task performance are helpful in this regard. When used prospectively, the operator or subject matter expert should be aided in creating a useful representation of the task as well as the system and operating environment which form the context of the task that is to be rated. In this latter case, the ratings of workload are made to descriptions of tasks and events that have not yet been personally experienced by the individual making the ratings (see Eggleston & Quinn, 1984).

REFERENCES:

Eggleston, R.G., & Quinn, T.J. (1984). A preliminary evaluation of a projective workload assessment procedure. In *Proceedings of the Human Factors Society 28nd Annual Meeting* (pp. 695-699). Santa Monica, CA: Human Factors Society.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock, & N. Meshkati (Eds.), *Human mental workload.* Amsterdam: Elsevier.

NASA-Ames Research Center, Human Performance Group (1986, Feb). *Collecting NASA workload ratings: A paper-and-pencil package* (Version 2.1). Moffet Field, CA: NASA-Ames Research Center.

Byers, J.C., Bittner, A.C., Jr. and Hill, S.G. (1989). Traditional and raw Task Load Index (TLX) correlations: Are paired comparisons necessary? In *Advances in Industrial Ergonomics and Safety*, Vol. 1. London: Taylor and Francis.

AVAILABILITY:

Human Performance Group
National Aeronautics and Space Administration
Ames Research Center
Moffet Field, CA 94035

- Pencil and paper version
- Computer version

Tr/Ta Task Analysis

DESCRIPTION: Task analysis methods seek to produce operator performance requirements as a function of fixed increments of time defined against a scenario background. These methods have a long history (Drury et al., 1987) and are the most commonly used of all analytical tools for predicting workload. General mission requirements are systematically decomposed into mission segments, functions, operator tasks, and detailed operator task element requirements. The result of the analysis is an operator activity profile as a function of mission time and segments, essentially a time-based analysis of performance requirements. In this context, workload is defined as time stress, and expressed as the ratio of the time required to perform a task (Tr) over the time available to perform the task (Ta).

SENSITIVITY: Individual operator or crew workload at the task level but only in terms of the time stress aspect of workload. Best utilized as an initial coarse filter to identify gross design deficiencies and for cases in which the time available for a task are well defined.

DIAGNOSTICITY: Limited to identifying general functional limitations where demands exceed operator capacity to respond within some time frame. Diagnosticity can be improved in the analysis partitions tasks into components relevant to sensory and motor channels (e.g., eye, ear, hand, and foot) and types of cognitive loads imposed upon the operator (e.g., target detection versus target identification).

INTRUSION:

IMPLEMENTATION REQUIREMENTS:

Data Collection: Detailed task analysis, to include task performance times, required for each mission.

Operator Training: n/a

OPERATOR ACCEPTANCE: n/a

SAFETY: n/a

PHYSICAL SPACE REQUIRED: n/a

PORTABILITY: n/a

INTEGRATION INTO SYSTEM: n/a

RESTRICTIONS: n/a

RELATIVE COST OF USE:

Testing time: n/a

Equipment: n/a

Setup and support:

Data analysis:

COMMENTS: Many variations of the time-based task analysis methods exists. A review of the different techniques may be found in Meister (1985). Some recent applications of time-based procedures include those of Edwards, Curnow, and Ostrand (1977) using the workload assessment model (WAM), and Linton, Jahns, and Chatelier (1977) using a variant of WAM, the statistical workload assessment model (SWAM).

REFERENCES:

Drury, C.G., Paramore, B., Van Cott, H.P., Grey, S.M., & Corlett, E.N. (1987). Task Analysis. In G. Salvendy (Ed.), *Handbook of human factors*. New York: Wiley.

Edwards, R., Curnow, R., & Ostrand, R. (1977). Workload assessment model (WAM) user's manual (Report D180-20247-3). Seattle, WA: Boeing Aerospace Co.

Linton, P.M., Jahns, D.W., & Chatelier, P.R. (1977). Operator workload assessment model: An evaluation of a VF/VA-V/STOL system. In *Proceedings of the AGARD Conference on Methods to Assess Workload* (AGARD-CP-216, pp. A12-1 - A12-11).

Meister, D. (1985). *Behavioral analysis and measurement methods.* New York: Wiley.

Type 1 Primary Measures

DESCRIPTION: Type 1 measures of primary task performance are indices of system+operator performance. Typically, they include measures of human tracking errors or measures of system performance. However, measures of system performance such as engine thrust, RPM, and movement of control surfaces could be classified as a Type 1 measurement, since changes in thrust, for example, reflect operator activities plus system lags. Similarly, any measure of effectiveness (MOE) for mission performance would ordinarily qualify as a Type 1 measure. Type 1 measures can also be called quality-related measures.

SENSITIVITY: Type 1 measures of system+operator are not often sensitive to subtle workload manipulations, but may indicate where overload has contributed to performance failure.

DIAGNOSTICITY:

INTRUSION:

IMPLEMENTATION REQUIREMENTS:

Data Collection:

Operator Training:

OPERATOR ACCEPTANCE:

SAFETY:

PHYSICAL SPACE REQUIRED:

PORTABILITY:

INTEGRATION INTO SYSTEM:

RESTRICTIONS:

RELATIVE COST OF USE:

Testing time:

Equipment:

Setup and support:

Data analysis:

Tools:

COMMENTS. Type 1 measures are important in system evaluation considerations and often are being measured anyway, outside of the context of workload.

REFERENCES:

Wierwille, W., Casali, J., Connor, S. and Rahimi, M. (1985). Evaluation of the sensitivity and intrusion of mental workload estimation techniques. In W. Roner (Ed.), *Advances in Man-Machine Systems Research*, Vol.2. (pp. 51-127). Greenwich, CT: J.A.I. Press.

AVAILABILITY:

FUTURE DEVELOPMENTS:

Type 2 Primary Measures

DESCRIPTION: Type 2 measures or primary task performance are those which assess the nature of operator performance directly. The measurement may be directed at quantify, frequency, or quality criteria of operator performance. Type 2 measures may also be directed toward detecting the fine structure of operator performance. In general, the category includes such measures as: (a) control movements per second in a psychomotor task, (b) response times in a perceptual or cognitive task, (c) errors of omission, (d) errors of commission, or (e) communications response times in a communications task (Wierwille et al., 1985).

The very reason Type 1 measures are insensitive is also the reason Type 2 measures are sensitive: As the operator copes with workload and under increasing load marshalls greater resources to hold Type 1 performance constant, the operator may perform differently and patterns of performance may change and fine structure tends to shift. Type 2, also called strategy-related or fine structure, measures are valuable because these shifts may provide evidence of a change in operator workload and hence provide a means to assess workload levels.

SENSITIVITY: Type 2 measures of the operator generally show effects on relevant dimensions.

DIAGNOSTICITY: May be more diagnostic than Type 1 measures as Type 2 measures may show shifts of operator strategies which may be useful in identifying sources of operator loading.

INTRUSION:

IMPLEMENTATION REQUIREMENTS:

Data Collection:

Operator Training:

OPERATOR ACCEPTANCE:

SAFETY:

PHYSICAL SPACE REQUIRED:

PORTABILITY:

INTEGRATION INTO SYSTEM:

RESTRICTIONS:

RELATIVE COST OF USE:

Testing time:
Equipment:
Setup and support:
Data analysis:
Tools:

COMMENTS:

REFERENCES:

Wierwille, W., Casali, J., Connor, S. and Rahimi, M. (1985). Evaluation of the sensitivity and intrusion of mental workload estimation techniques. In W. Roner (Ed.), *Advances in Man-Machine Systems Research*, Vol.2. (pp. 51-127). Greenwich, CT: J.A.I. Press.

AVAILABILITY:

FUTURE DEVELOPMENTS:

Unstructured Interviews

DESCRIPTION: Unstructured interviews are those in which questions are asked based strictly on observation or in response to what the interviewee has said. Unstructured discussions should be used whenever possible to obtain the subtle, detailed information that might not be obtained from rating scales.

SENSITIVITY: Variable.

DIAGNOSTICITY: Detailed information that might not otherwise be obtained can be drawn from interviews.

INTRUSION: Most appropriate to interview after the test session or operational sequence to obtain the most information and reflections of the operator.

IMPLEMENTATION REQUIREMENTS:

Data Collection: An interviewer who is familiar with the system/operation under study is necessary so that subtle information regarding OWL can be drawn from the interviewee. Video or audio recording, or paper and pencil transcription.

Operator Training: Minimal, although operators may become more comfortable with expressing opinions as they become more experienced in being interviewed.

OPERATOR ACCEPTANCE: Interviewing is a well established method of obtaining subjective information. Cooperation can generally be expected to be good, but there may be individuals who are uncomfortable and will be uncooperative. In addition, there can be the problem of individuals who are unwilling to give negative reports. Practitioners should be aware of possibly misleading information obtained through interviews.

RELATIVE COST OF USE: The highest cost is the time invested by the interviewer.

Testing time: Variable, although less than 30 minutes is recommended.

Equipment: Recording equipment, if used.

Setup and support: Minimal.

Data analysis: Qualitative.

COMMENTS:

REFERENCES:

Dyer, R., Matthews, J., Wright, C., & Yudowitch, K. (1976). *Questionnaire Construction Manual* (TCATA DAHC-19-74-C-0032). Ft. Hood, TX: ARI.

U.S. Army Test and Evaluation Command (1976). *Questionnaire and interview design, Subjective testing techniques* (TECOM Pam 602-1, Vol. I). Aberdeen Proving Ground: USATECOM.

Meister, D. (1985). *Behavioral analysis and measurement methods*. New York: John Wiley and Sons.

Ericsson, K. A. (1984). *Protocol Analysis*. Cambridge, MA: MIT Press.

Zachary/Zaklad Cognitive Analysis

DESCRIPTION: This is a methodology for detailed assessment of operator cognitive workload during a tactical mission. There are two phases -- (1) the construction of a detailed cognitive mission task timeline and (2) workload analysis based on that timeline. (1) cognitive timeline -- a specific mission scenario is required. A cognitive goal-based analysis of mission prosecution is constructed -- iteratively with input from both operational SMEs and cognitive scientists. Strategies are outlined for performing the whole range of operator functions, from button-pushing to making complex sensor management decisions. The mission scenario is decomposed into appropriate time units and the cognitive model applied to each segment, resulting in a detailed timeline of operator activity, display status, and external events. (2) workload analysis -- a detailed channel-specific rating scale with 13 subscales including 5 cognitive is used and tailored to the particular application. A different group of SMEs then rates the level of each type of workload during each mission segment of the timeline developed in step (1).

SENSITIVITY:

DIAGNOSTICITY: Very high, 13 different workload subscales, mission-tailored, and particularly detailed in cognitive area

INTRUSION: Minimal, conducted completely outside of mission activities.

IMPLEMENTATION REQUIREMENTS: Low to moderate

Data Collection: Extended interviews and subjective rating sessions with designated SMEs

Operator Training: Two groups of SMEs are required, minimal training required on both the timeline and ratings steps,

OPERATOR ACCEPTANCE: Usually very high

RELATIVE COST OF USE: Low, main labor cost is training the experimenters to understand the operational environment.

Testing time: step 1 -- several days, step 2 -- several hours

Equipment: None

Setup and support: Private room required to conduct sessions

Data analysis: Variable, can be minimal if no statistical techniques are used, but can be more extensive -- 13 subscales X #mission segments X #subjects.

Tools: Strictly pencil and paper

COMMENTS: Two main problems. First, Z2 relies on subjective judgments, albeit in a systematic manner. Second, Z2 has a very minimal track record and documentation. It was applied to two P-3 TACCO missions (Zaklad, et al 1983) with some success, and to a single F/A-18 mission (Zachary et al, 1987; Zaklad et al, 1987) in which the timeline was constructed but the workload analysis was not completed (for non-technical reasons).

REFERENCES:

Zaklad, A.L., Deimler, J.D., Iavecchia, H., and Stokes, J. (1982). *Multisensor correlation and TACCO workload in representative ASW and ASUW environments.* Analytics TR 1753A.

Zachary, W., Zaklad, A., & Davis, D. (1987). A cognitive approach to multisensor correlation in an advanced tactical environment. *Proceedings of the 1987 Tri-Service Data Fusion Symposium*, Johns Hopkins University, Laurel, MD.

Zaklad, A., Zachary, W., & Davis, D. (1987). A cognitive model of multisensor correlation in an advanced aircraft environment. *Proceedings of the Fourth Midcentral Ergonomics/Human Factors Conference*, Urbana, IL.

AVAILABILITY:

Naval Air Development Center
Warminster, PA 18974

FUTURE DEVELOPMENTS: Validation of the method on a variety of mission contexts.

APPENDIX B

OPERATOR WORKLOAD KNOWLEDGE-BASED EXPERT SYSTEM TOOL (OWLKNEST) SURVEY

OPERATOR WORKLOAD KNOWLEDGE-BASED EXPERT SYSTEM TOOL (OWLKNEST) SURVEY

Thank you for taking the time to fill out the attached questionnaire. This is the first version of OWLKNEST that we have distributed. We feel that you, as a potential user of the tool, are the best source to evaluate OWLKNEST. Your suggestions and comments will be carefully considered.

The more specific and detailed your responses, the easier it will be for us to discern the problem and incorporate your suggestions in the next revision of OWLKNEST. If there is not enough space for your comments, please feel free to continue on the back of the questionnaire pages or attach additional pages.

We would like to contact you if we had some questions about any of your responses. Please fill out the following:

NAME ________________________________

ADDRESS ______________________________

____

____

PHONE ________________________________

Please return the completed survey to:

Richard E. Christ
US Army Research Institute Field Unit
P O. Box 6057
Attn: PERI-SB
Ft Bliss, TX 79906-0057

OPERATOR WORKLOAD KNOWLEDGE-BASED EXPERT SYSTEM TOOL (OWLKNEST) QUESTIONNAIRE

INSTALLATION AND START-UP

Did you encounter any difficulty installing OWLKNEST and gettingthe program to run? If yes, please explain.

APPLICATIONS

Briefly describe the applications that you tried with OWLKNEST. Explain the basic framework for the hypothetical situation or the actual study you plan to conduct or have conducted. (See the Example Section in HOOT).

OWLKNEST DIALOGUE: QUESTIONS AND RESPONSE ALTERNATIVES

1. Which, if any, of the questions were not clear?

2. Were you able to find suitable response(s) among the alternatives listed for each question? If not, then please specify.

3. How often did you use the <?> DETAILS query to clarify your understanding of the question and/or alternatives?

   never |___|___|___|___|___|___| very often

4. How useful was the information contained in the <?> DETAILS section?

5. How clear was the output (i.e., the selected categories and recommended techniques)?

FEATURES OF OWLKNEST

1. Did you try the <C> CHANGE and RERUN option? If yes, please comment on the utility of this option.

2. How informative (useful) were the rules displayed with the WHY option?

3. How useful was the <H> HELP feature which explained the stage of the program and what response(s) were required?

4. How easy was it to save your output to a data file?

5. How easy was it to obtain a printout of your results?

6. Any problems exiting the program when desired? If yes, then please describe.

CONTENT

1. How many of the suggested techniques would adequately address issues for your particular workload study(ies) (i.e., based on your inputs)?

   none |___|___|___|___|___|___| all

2. Are there any workload techniques that should be added or removed to the OWLKNEST knowledge base? If so, what are they and why do you recommend the change?

3. Were there any vital issues which were not addressed by the OWLKNEST questions? If so, then please explain.

4. How useful was the information on the techniques?

OVERALL EVALUATION

1. How useful were the ratings? Please explain.

2. How satisfied are you with OWLKNEST as a workload guidance-providing tool?

3. How satisfied are you with the functioning of the expert system shell (EXSYS) (e.g., how responses are input, how the WHY and HELP commands function, etc.)? What features did you like or dislike?

4. How satisfied are you with the OWL handbook, HOOT? What features did you like or dislike?

5. Which of the following best describes the level of your understanding about how OWLKNEST utilizes your input to arrive at the recommended techniques?

____ Don't understand it at all and don't feel it is necessary to understand it in order to effectively utilize OWLKNEST and have confidence in the results

____ Don't understand it and would like to feel more comfortable about the reasoning and logic behind the recommendations

____ Have grasped the fundamentals

____ Have grasped the fundamentals and would like to understand more

____ Have a full understanding

General comments about OWLKNEST:

BACKGROUND

1. What is the extent of your knowledge in the area of operator workload?

minimal |___|___|___|___|___|___|___| extensive

2. What is your level of computer experience?

novice |___|___|___|___|___|___|___| expert

3. What is your level of experience with expert systems?

novice |___|___|___|___|___|___|___| expert

THANK YOU FOR YOUR FEEDBACK!